Evaluating the Information Architecture of Digital Museums

by

Megan Sellmer

A thesis submitted in partial fulfillment of the requirements for the degrees of

Master of Arts in Humanities Computing
and
Master of Library and Information Studies

Humanities Computing / Library and Information Studies
University of Alberta

Abstract

Digital museums need to prioritize the information architecture found in their website design in order to optimize the structure and organization of their collection for users. This study used the usability inspection method, heuristic evaluation, to examine nine select digital museums. The heuristics were created during a content analysis study of general knowledge information architecture books. The information architecture principles found during the heuristic evaluation were compared and summarized in order to create a list of criteria that select digital museums can use to facilitate interoperability and consistency between websites.

Acknowledgements

I would like to thank and acknowlegde the guidence of my thesis advisors, Geoffrey Rockwell and Ali Shiri. Their persistance and input was invaluable to me throughout this process. I would also like to thank Tami Oliphant and Sean Gouglas for agreeing to chair and sit on my committee. A special thank you goes to my family and friends for supporting and encouraging me throughout this process. You are an amazing group of people and I could not have done this without you.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Introduction

Digital museums are online environments that use different technologies (like 3D graphics and multimedia) to present collections of objects with contextual information in order to create an experience for users (Foo, 2008; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009; Zhou et al., 2012). Providing access to content and contextual information is important for the success of these sites, which is why the design of their information architecture (IA) needs to be well constructed. Information architecture works in digital environments by structuring and organizing information, so that users can find and understand the information presented online (Rosenfeld et al., 2015). A list of information architecture criteria, derived from IA components found in live websites, could improve how select digital museums design their information architecture as well as promote interoperability between those sites and consistency in how users access information.

A heuristic evaluation of nine select digital museums revealed the current uses of information architecture in live sites, which theoretically works well for users (and have likely undergone user testing prior to being published). The heuristics used in the examination will be created during a qualitative content analysis study of information architecture literature. By compiling a list of the information architecture principles found in those sites and then comparing and contrasting them against each other one can see what IA principles select digital museums have determined necessary for the structure and labeling of the digital museums. This list of information architecture principles could be considered criteria or IA components. Select smaller websites and currently developing digital museum websites could then use this list to make sure that their IA design is comparable to other digital museum sites (their graphic design and content can obviously vary), thus promoting consistency and interoperability.

Research Questions

The questions that this research study will set out to answer are:

1. Can the current information architecture design in digital museums be condensed into a list of criteria to promote consistency between like websites and ease of use?
2. What are the information architecture principles present in the select information architecture literature?
3. How do select digital museums use information architecture principles in their live website design?
4. From the results of the study, what information architecture criteria should digital museums have?

To determine the information architecture structure found in live digital museums, a heuristic evaluation of nine select digital museums will be conducted (three of each type of digital museum, for more information see below). The heuristics (a set of guidelines and definitions used during the website evaluation) will be created during a content analysis study of four general knowledge information architecture books. These books will contain most, if not all, the principles found in information architecture because they cover such a broad-range of IA information. Through multiple examinations of the digital museums, the heuristic evaluation will allow the researcher to learn exactly what information architecture components each digital museum uses, as well as if there are any information architecture bugs/issues in the sites. The findings will create the list of criteria that select digital museums can use when designing their information architecture and possible solutions for the bugs/issues found in the digital museums examined.

Background

It is difficult to define the term digital museum because there are thousands of digital museums on the web, each with different collections that can contain anything from Renaissance paintings to carpenter tools. Additionally, the definitions can vary according to the creator's background and interests (a computer scientist has a different definition than a museum curator). There are, however, key components of a digital museum that have general consensus, enough to build a broad definition that can then be narrowed down into different classifications. Digital museums are online environments that use different technologies (like 3D graphics and multimedia) to present collections of objects and contextual information in order to create an experience for users (Foo, 2008; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009; Zhou et al., 2012). Information architecture components facilitate access to information, which is why it is so important to get it right.

We can further categorize digital museums into "types." Here are the definitions, summarized from multiple sources, of the three different types of digital museums – brochure, content, and learning.

- **Brochure** digital museums are the online representation of a real museum. It follows the mission statement and goals of the organization, and focuses on marketing the museums and its contents to potential visitors.
- **Content** digital museums are the online catalogue of a museum or a collection of objects oriented around a topic, time period, resource, etc.
- **Learning** digital museums emphasizes education for their audience, about the digital objects in their collection. Often the learning museums will link to additional sources to motivate users to research an object of interest.

(Styliani et al., 2009; Schweibenz, 2004; McDonald, 2005; Piacente, 1996; Teather 1998)

This study will examine three examples of each type of digital museum.

Digital museums began in the 1990s, with the introduction of the World Wide Web. Museums began to digitize materials on CD-ROMs for preservation and cataloguing, so once user-friendly browsers like Mosaic became available, museums and individuals began to upload those images to digital or virtual museums (Bianchini, 2016; Mannoni, 1996). Bowen (1999) describes an early digital museum, *The Museums of the History of Science* (part of the University of Oxford) entitled "The Measurers: a Flemish Image of Mathematics in the Sixteenth Century."[1] Originally published in 1995, the design has not been updated since so it offers a glimpse of how digital museums were first structured (Bowen, 1999). Another early example was created by the French Ministry of Culture, which began an elaborate digitization project in the mid 1990s; their goal was to provide online access to everyone in order to protect objects that were in poor

[1] This exhibit is still accessible at www.mhs.ox.ac.uk/measurer/text/contents.htm

physical condition (Mannoni, 1997). As technology progressed and people became more and more familiar with it, the number of digital museums increased substantially. This introduced a greater need for information architecture.

Technology (like 3D graphics and image enhancement) helped transform how users view and understand digital objects. However, in the rush to use these technologies, information architects need to keep in mind how they affect the usability of a website (Knell, 2003; Morville et al., 2016; Styliani et al., 2009; Tam & Robertson, 2002; Wyman et al., 2011). Information architecture should structure and present those technologies in a way that users can access and understand. Studies have shown that following best practices when designing the information architecture of a website helps create a trust between users and the website (we are comfortable with the familiar)(Riley-Huff, 2012; Parandjuk, 2010). It is not just new technologies that benefit from information architecture because IA organizes the content and accessibility of the entire website, which is why it must be updated and reassessed as the website grows or improves.

Richard Saul Wurman (1997) coined the phrase information architecture, but his work also discusses *Information Architects* and their roles in planning and designing a website. He describes them as people who discover and organize data in order to provide paths so that users can access that information (Wurman, 1997). Information architects can improve the user experience "… by recognizing the importance of perspective, by striving to understand the intended audiences through user research and testing, and by providing multiple navigation pathways…" (Morville & Rosenfeld, 2006, 57). An information architect must consider the components of information architecture (organization, navigation, search, labeling, and vocabulary systems). That is why these are the main categories for the content analysis study.

A content analysis will be used to examine the information architecture literature. This is the formal evaluation of text according to a coding scheme (created by the researcher), in order to understand a deeper meaning in the text and code it for further analysis (Mayring, 2000; Schreier, 2012; Schreier, 2014). The information architecture principles established by the content analysis will be used as the heuristics during the evaluation of the digital museums.

A heuristic evaluation allows the researcher to discover how digital museums use information architecture components within their sites, and if there are any issues associated with them. In a heuristics evaluation, three to five expert evaluators examine a website looking for issues or bugs using a list of guidelines and/or principles to search with (the heuristics) (Nielsen, 1994a; Nielsen & Molich, 1990; Novick & Hollingsed, 2007; Sauro, 2010). For this study there will be only one expert evaluator, who will examine the nine digital museums using the IA heuristics. The evaluator will go over the websites three times, once to become familiar with the site, once to perform the heuristic evaluation, and finally once to confirm the findings. This is to ensure that there are no false positives – problems that are not actually issues, just the evaluator's opinion (Cockton & Woolrych, 2001; Sauro, 2012a; Sauro, 2012b; Sauro, 2016). Comparing all the IA principles found in the heuristic evaluation will create the information architecture criteria that select digital museums can use as a guide for their IA design. These will be summarized and then presented in list form for clarity. The bugs/issue found on the site will be researched and solutions suggested to fix them.

For a list of definitions related to this project see the glossary, Appendix A.

Chapter Summary

This study will look to create information architecture criteria from current IA principles found in the nine select digital museums. This will promote consistency and interoperability between select digital museums and organize the content so users can easily access information. This will be accomplished in two stages, (a) a content analysis of information architecture literature to discover the key IA principles, and (b) a heuristic evaluation of nine digital museums using the list of IA principles as the heuristics. Data collection, analysis, and research findings will be presented in great detail, so that the readers may judge the study for themselves.

Relevant information about both digital museums and information architecture will be discussed in the literature review. Topics that will be covered include history, advantages and disadvantages, and digital museum audiences.

Chapter 2: Literature Review

Introduction

This study is made up of two main components, digital museums and information architecture. The following literature review defines and provides background information about these two areas of study.

Digital museum is a broad term that can include many different types of websites; an overview of different definitions will be presented to provide an overall picture of how the literature defines this term. This includes a closer look at the three types of digital museums found in the literature, which the digital museums evaluated in this study can be classified into. The history of digital museums will be outlined, with a closer look at the two branches of history that merged to create digital museums –museums and modern technology. There will also be an overview of the evolution of digital museums from CD-ROMS to early examples on the web. The advantages and disadvantages that digital museums offer users and physical museums will be provided for more information about how digital museums impact users. This section concludes with a look at the different types of visitors that use digital museums (and why they use them).

The information architecture (IA) of websites is a vital component for the structure and organization of content. To provide an overview of this field, a definition of information architecture will be discussed along with additional information, such as IA website classifications (organization, navigation, labeling, search, and vocabulary systems) and why this field of study is so important. An outline of a website's contextual framework will be provided because it is an important consideration when evaluating and designing information architecture. This section will also include a summarized history of information architecture in order to contextualize the field of study for the reader.

Digital Museums

Definition

There are many different definitions for digital museum, even the term changes from source to source (e.g., virtual museum, online museum, virtual exhibit, cyber museum, etc.). The reason for this inconsistency is because creating digital museum involves different specialists (programmer, museum curator, usability specialist, etc.) and different organizations (museums, businesses, charities, etc.), all of who have their own definitions and terms for these websites. Below the author will examine some of these definitions and outline the three different types of digital museums.

Digital museums can be described in many ways; however, there are some themes that consistently show up in the definitions. A basic definition (the one that will be used for this study) is an online collection of digital objects, presented to the user with different technologies that are traditionally associated with a physical museum but can be attached to different online organizations (Foo, 2008; Karp, 2004b; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009). This broad definition can represent the three different types of digital museums (brochure, learning and content – discussed in further detail below). They are created using current technologies and information architecture to present an engaging, learning focused, and user-centered websites. The technologies that digital museums use include multimedia, virtual reality, 3D graphics, and much more (Styliani et al., 2009; Wang et al., 2014; Zhou et al., 2012). This combination of digital collections with evolving technologies makes digital museums a dynamic field of study, which would benefit from standardization (beginning with the list of criteria presented in this study).

Definitions of digital museums also focus on linking information between the collection and contextual information. The basic tenet of digital museums, according to Glen Hoptman (1992), is the connectedness of information, between objects and contextual descriptions, which turn digital museums into education resources. Digital museums allow users to learn with contextual information, more than what the physical museum can provide (either within an exhibition or for an object), guiding users to more information or different points of view. Werner Schweibenz (1998) also emphasized contextual links when defining digital museums. He said that digital or virtual museums connect visitors with objects and information through contextual links to related/similar content (for example comparing works of art by the same artist or artwork completed by artists from the same period) (Schweibenz, 1998). This connectedness is directly related to a digital museum's information architecture, as contextual links are part of IA navigation systems. Hoptman and Schweibenz focus on the importance and ability that digital museums have of providing additional information to users.

When defining digital museums, some authors do so by comparing them to digital libraries. Jane Barton (2005) defined digital museums as a collection of digital objects with metadata, which allows content retrieval and display (like a digital library). She goes on to discuss the difference between digital museums and digital libraries and how they can be integrated to become cultural institutions, which would be easier to do if they used the same metadata standards (Barton, 2005). This definition is closely related to the definition of digital libraries (but with cultural heritage instead of books). Dagny Stuedahl (2007) also made this comparison, defining digital museums as digital libraries that primarily focus on cultural heritage. This is connected to his theory of convergence, with libraries, archives, and museums coming together to create vast interconnected information repositories with up to date technology (Stuedahl, 2007). Convergence of digital media can bring collections together, but the technologies and standards that the different organizations use, make this interconnectedness difficult (Stuedahl, 2007).

A review of definitions for digital (or in this case, virtual) museums would not be complete without including the definition from the *Encyclopædia Britannica* (2017). It defines digital (or virtual) museums as

> "... a collection of digitally recorded images, sound files, text documents, and other data of historical, scientific, or cultural interest that are accessed through electronic media. A virtual museum does not house actual object and therefore lacks the permanence and unique qualities of a museum in the institutional definition of the term." (Encyclopædia Britannica, 2017, n.p)

The *Encyclopædia Britannica* (2017) mentions that the definition of virtual museums do not follow the "institutional definition" for museums, a reference to the International Council of Museums (ICOM), which states "A museum is a non-profit, permanent institution…"

(International Council of Museums, 2007, Museum Definition). They have not updated this definition in ten years. Considering the advancements in technology, it should be.

The definition of digital museums used for this thesis is similar to those discussed above, online environments that use different technologies (like 3D graphics and multimedia) to present collections of objects with contextual information in order to create an experience for users (Foo, 2008; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009; Zhou et al., 2012). A list of criteria for digital museums can create connections between other heritage institutions to promote consistency and interoperability. This broad definition can fit many different websites, but because digital museums can contain different content and different technology the definition needs to be as inclusive as possible. The nine select digital museums evaluated in this study can all be described using this definition.

Types of Digital Museums

The literature referenced three different types of digital museums, brochure, content, and learning. Here is a more detailed explanation of the three digital museum types, summarized from multiple sources:

- Brochure digital museums are the online representation and marketing tools of physical museums, containing visitor information like hours, location and events. Physical museums use the web to present their collection, and exhibitions, promote themselves and supplement the museum experience for their target audience. This type of digital museum has evolved since the 1990s when it generally consisted of just location and operating hours information. Now digital museums are using technology and digital objects to create an experience for users in order to entice them to go and see the real thing.
- Content digital museums are the online catalogue of a museum or a collection of objects oriented around a topic, time period, resource, etc. These websites aim to make large collections of art available to the public, often presenting content in an object-oriented way with contextual information. It can have the same features as both the brochure and learning museums, but it is often simpler (a database with metadata).

- Learning digital museums are websites that organize information for specific audiences with links to educational resources. They provide contextual information to motivate users to further study the items (and to encourage users to return). Contextual information can include links to related objects within the digital museums or historical/reference information on other websites. They can be games, stories, articles/books, etc. (generally dependent on the target audience).

(Antonaci & Ott, 2014; Barton, 2005; Chae & Kim, 2011; Cody, 1997; Foo, 2008; McDonald, 2005; Piacente, 1996; Schweibenz, 2004; Styliani et al., 2009; Teather, 1998).

History

Digital museums are the product of two historical timelines – museums and modern technology (computers and the Internet). Both will be examined below to see how the timelines connected to create digital museums.

Museums

The word museum comes from the Greek word "Mouseion" translated to temple of the muses, the Greek goddesses of the arts (both literary and scientific) (Alexander, Alexander & Decker, 2017; Findlen, 1989; Maranto, 2015). These temples were filled with offerings to the goddesses, including sculpture, mosaics, gold, gemstones, scientific instruments, and poems (Alexander, Alexander & Decker, 2017; Findlen, 1989; Maranto, 2015). While Greece is where the term "museum" originated, the first known museum (or curated collection available for viewing by the upper-class) was in Mesopotamia, 530 BC (around the area of Iraq) created by the princess Ennigaldi-Nanna in her home; artifacts were laid out in rows with clay labels (written in three languages) (Maranto, 2015). Fast-forwarding a few years, the Romans displayed public art in forums, public baths, gardens, etc. but these were not available to all members of Roman society (despite the term "public"), just the upper and middle classes (Alexander, Alexander & Decker, 2017). This exclusion of the lower classes would be a common theme found in museum history until the 20th century.

Roman Generals and politicians collected objects from conquered tribes during the Roman expansion across Europe; this is the first version of what would become known as cabinets of curiosities (Maranto, 2015). Cabinets of curiosities (as we know them) became popular between the 1400 and 1600 hundreds; they housed collections of natural history and artistic works in anything from a small box to multiple rooms (Amsel-Arieli, 2012; Grice, 2015). The wealthier the collector the more elaborate the cabinets of curiosities were, these could include gems, minerals, taxidermy, skeletons (animal and humans), first peoples art, and plants amongst other things – these were collected by explores and merchants who travel to other lands and brought back these treasures to keep or sell (Amsel-Arieli, 2012; Grice, 2015). The most famous cabinets of curiosities in history include:

- Ole Worm (a Danish physician, artist philosopher, etc. 1588-1655) who had a large collection that mainly included animals. He catalogued these items by drawing and labeling them. His collection included a shark, squid, the skull of a male narwhal, etc.
- John Tradescant the Elder (an English royal gardener, 1570-1638) called his cabinet of curiosity the "Ark," which was filled with oddities and botanicals collected in his travels. His son also added to this collection, which they organized into sub collections like portraits, animals, plants, religious relics, etc. Eventually they opened up their cabinet of curiosity for public (the upper class) viewing.
- Hans Sloane (Irish physician, 1660-1753) had one of the largest cabinets of curiosities in England. He bought (or was given) other collector's cabinets of curiosities and meticulously catalogued each item (comparing, identifying, and classifying). His collection of 71,000 objects formed the foundation for the British Museum (1881).

(Amsel-Arieli, 2012; Hagen, 1876; Grice, 2015)

Collections of natural and artistic objects occurred throughout history, Alexander the Great and Emperor Augustus both collected objects of interest. However, the collection of natural history became far less popular after the death of Christ (and the birth of Christianity) because this period emphasized religious piety over the appreciation of art and science, it was only the monasteries that collected works of the past for the preservation of knowledge (like religious artifacts like relics, manuscripts, etc.) (Alexander, Alexander & Decker, 2017; Hagen, 1876). It was the Renaissance that made collecting fashionable again. As cabinets of curiosities moved from small to large, private to public, these collections needed to be preserved – boards of trustees were appointed (due to wealth and not necessarily knowledge), and then curators for the collections (Hagen, 1867). The first public museum was the Ashmolean Museum at Oxford, opening in 1683, when Elias Ashmole donated his cabinet of curiosity collection, which contained scientific, magical, and natural history objects to Oxford University (Alexander, Alexander & Decker, 2017; Findlen, 1989). Another collector was Phineus T. Barnum, who bought cabinets of curiosities from Americans and Europeans in the 1840s. He opened a public museum in New York that contained not only objects from the purchased collections, but also unique individuals (for example, Siamese twins) live animals, technology, wax figures, and much more (Maranto, 2015).

The museums we think of when we hear that term today are called survey museums. These display a broad array of art history – in both the medium of the objects (paintings, sculpture, textiles, pottery, etc.) and the location or culture they are from (China, First Peoples, Renaissance, prehistoric, etc.) (Duncan & Wallach, 2004). The 19th and 20th century saw an increase in the popularity of these museums, especially when they started allowing all members of the public to view the collection. A few of the most influential of those would be The Louvre, The Metropolitan Museum of Art and The Smithsonian Institute (Duncan & Wallach, 2004).

- In 1793, the Palace of the Louvre was opened to the public; they could view the collection three days a week, the first survey museum to do so (Alexander, Alexander & Decker, 2017; Duncan & Wallach, 2004; Maranto, 2015).
- Englishman James Smithson bequeathed a sum of money for the creation of the Smithsonian Institute (Alexander, Alexander & Decker, 2017; “Our History,” 2017). The donation was accepted in 1836, but Congress debated creating the museums until the bill was passed in 1846, signed by President James K. Polk (“Our History,” 2017).
- Metropolitan Museum of Art was incorporated in 1870, but it opened its doors in 1880 (Alexander, Alexander & Decker, 2017, “History of the Museum,” 2017). This Museum partially came about because of national pride, museums were prestigious cultural institutions found across Europe (even in smaller cities), and New York wanted to have one of their own. (Duncan & Wallach, 2004). The very first item collected by this museum, bought in 1870, was a Roman sarcophagus (“History of the Museum,” 2017).

Technology

The history of the computer and Internet can be seen both in individual and collaborative contributions. Throughout the last three centuries there have been people who have worked to develop the computer, including Ada Lovelace, Alan Turing, and Vannevar Bush. The Internet was much more of a collaborative project, brilliant people working to create an interconnected future. These histories will be outlined below.

Charles Babbage (1791-1871) was the son of a banker and a Cambridge graduate; he wrote books about a variety of subjects including science, insurance, and machinery (Wilkes, 2002). He was one of the first people to discuss what would become a precursor to the computer. Charles Babbage (1864/2011) described a device that would solve mathematical calculations, called the Analytic Engine (though it was never built). This machine would be made up of two parts, (a) where you input the calculation using punch cards, and (b) where you receive the results (Babbage, 1864/2011). To use the Analytic Engine you needed operation cards (to represent what the user wanted it to do), variable punch cards (to input the variables in the calculation), and number cards (to specify the numbers used), which could be strung together to form complex queries (the punch cards were based on Jacquard's loom) (Babbage, 1864/2011). Babbage (1864/2011) concluded that the biggest difficulty was not actually building the machine, but making it work for the different formulas that were being introduced in the field of mathematics.

Augusta Ada King, the Countess of Lovelace (1815-1852) was the daughter of the poet Lord Byron and mathematician Anna Isabelle Milbank (Aiello, 2016; Coe & Ferworn, 2016; Essinger, 2014). She was a close friend of Charles Babbage and is considered by most to be the first computer programmer for the Analytic Engine (see above) (Aiello, 2016; Essinger, 2014). When Ada translated a French article by Luigi Federico Menabrea about the Analytic Engine, Charles Babbage encouraged her to add her own notes to the article, and this is where we see her greatest impact on the future of the computer (Babbage, 1864/2011; Lovelace, 1842). Ada understood that the Analytical Engine could do more than just solve mathematical formulas; the pattern recognition could be used for digitizing music or working with biology models (Aiello, 2016; Essinger, 2014; Lovelace, 1842). What could be considered ‘computer programming’ came in her last note (Note G) of this article, where she detailed how the Analytical Engine would solve the Bernoulli Numbers (how to design the punch cards, etc.) (Aiello, 2016; Coe & Ferworn, 2016; Essinger, 2014; Lovelace, 1842). Ada was a brilliant mind in a time that disvalued female contributions to science, it would have been interesting to see what else she would have come up with if she had not died from cancer at 36 (Aiello, 2016; Essinger, 2014).

Alan Turing is the founder of computer science (Bowen, 2017). A brilliant British mathematician, he conceptualized the precursor to the computer – the Automatic Computing Engine (ACE), which was built after he left the National Physical Laboratory in 1947. Turing is (possibly) most famous for his work in Bletchley Park and designing the bombe to break the German Enigma coding machine, which he did by identifying likely text and working backwards through the Enigma combinations (Bowen, 2017). Turing also predicted the eventual creation of artificial intelligence (AI) and devised a test to determine if the machine was an AI, it was called the imitation game (Bowen 2016; Bowen, 2017; Turing, 1950). In this test an evaluator asks questions to determine what they are examining, a machine or a person. These can be any questions, for example what is 38957182 x 86 – a machine will likely answer quickly and accurately but a human will probably have great difficulty in answering this without a calculator (the answer is 3350317652) (Turing, 1950). However, there needs to be more questions than just math because a computer can be programmed to answer incorrectly.

Vannevar Bush was an engineer who contributed to the advancement of technology in two ways, the differential analyzer and the memex. The differential analyzer was built at the Massachusetts Institute of Technology (MIT) in order to solve differential equations (the calculation between physical quantities and rates of change) (Bush, 1931). In the article, *The differential analyzer. A new machine for solving differential equations*, Vannevar Bush (1931) describes why the device was built and how to use it (to achieve accurate results). After World War II, Vannevar Bush (now the director of the Office of Scientific Research and Development in the United States) in his article *As We May Think* (1945) urged scientist to continue the collaboration they conducted to aid the war efforts, specifically he said that they needed to organize, transmit, and review each others knowledge. One way to do this Vannevar Bush (1945) proposed, was to create the memex machine (a type of personal computer), which would let users view files and books stored as microfilm on the machine. The memex would also let users link files and items together by adding the codes at the bottom of the file, much like hypertext links (Bush, 1945).

Just over ten years after Alan Turing and Vannevar Bush made major contributions to the creation of computers, the Internet began. The Russian launch of the Sputnik satellite in 1957 was what inadvertently launched the Internet. In response to that event, the US government created the Advanced Research Project Agency (ARPA) as part of the department of defense, which produced the Advanced Research Project Agency Network (ARPANET) the precursor to the Internet (Campbell-Kelly et al., 2014; Cohen-Almagor, 2011; Leiner et al., 1997; Leiner et al., 2009; Williams, 2017). Those sources describe ARPANET as a network of nodes that (eventually) spread across the globe, connecting research institutions. In 1961, Leonard Klienrock who worked at the MIT, introduced the idea of packet switching, which is when chunks of information (the packets) are separated and sent separately over the network only to be reassembled at its destination (this made transmissions faster and reduced delays of the messages) (Campbell-Kelly et al., 2014; Cohen-Almagor, 2011; Leiner et al., 1997; Leiner et al., 2009; Williams, 2017). ARPANET used packet switching when Bolt Beranek and Newman (BBN) launched it in 1969, the four different computer models connected together had Interface Message Processors (IMPs) so that they could communicate with each other and transmit files and messages (located at University of California at Los Angeles, Stanford Research Institute, University of California at Santa Barbra, and the University of Utah) (Campbell-Kelly et al., 2014; Cohen-Almagor, 2011; Leiner et al., 1997; Leiner et al., 2009; Williams, 2017).

It wasn’t until 1974 that Vint Cerf and Robert Kahn (often called the fathers of the Internet) would introduce the Transmission Control Protocol and Internet Protocol (TCP/IP) (Leiner et al., 1997; Leiner et al., 2009). TCP would improve on packet switching by setting rules for how data was organized into packets and how they were reassembled on their destination; IP defined rules for how the data packets were routed on their way to their destination (Cohen-Almagor, 2011; Leiner et al., 1997; Leiner et al., 2009; Williams, 2017).

ARPANET incorporated TCP/IP as it expanded across the globe (by 1983 all networks used it) (Cohen-Almagor, 2011; Williams, 2017). One of the key aspects of ARPANET was the idea of open architecture, which meant, "Each network could continue using its own protocols and data-transmission methods internally. There was no need for special accommodations to be connected to the Internet, there was no global control over the network, and all could join in" (Cohen-Almagor, 2011, 51). This is still present in todays Internet and is why we can connect to the Internet using any computer, with any operating system, anywhere in the world (well as long as the computer is new enough and you pay a fee, but it is not just limited to only university teachers and students, for example).

Other notable technological advancements can be seen in the modern development of the computer. IBM (International Business Machines Corporation) was the main producer of computers in the 1950s and 1960s; these were mainly sold to government and corporations (Madrigal, 2011). In 1980, IBM began using an operating system called Microsoft in their corporate/government machines; they would produce a personal computer with the Microsoft operating system a year later (Madrigal, 2011). Microsoft was of course created by Bill Gates and Paul Allen in 1975 and would be one of the most successful computer related companies ever (Campbell-Kelly et al., 2014). Also in 1975, Steve Jobs and Stephan Wozniak founded Apple Inc. (one of the other most successful technology companies) (Campbell-Kelly et al., 2014). These two companies would shape the personal computer for years to come.

In 1989, Tim Berners-Lee created the World Wide Web (working off ARPANET) (Cohen-Almagor, 2011; Berners-Lee et al., 1992; Williams, 2017). He defined the Uniformed Resource Locator (URL), Hypertext Transfer Protocol (HTTP), and HyperText Markup Language (HTML) (Cohen-Almagor, 2011; Berners-Lee et al., 1992; Williams, 2017). Tim Berners-Lee is a British physicist and software engineer who was working at CERN (the European Organization of Nuclear Research) when he introduced his idea of a "… system of international protocols called the World Wide Web" (Cohen-Almagor, 2011, 53; Williams, 2017). The World Wide Web lets users access the web in two ways, through search (text retrieval) and with hypertext links, the pilot project for WWW determined that those were the most important navigation tools for users (links between similar information) (Berners-Lee et al., 1992). Ted Nelson (1987) coined the actual word "hypertext" in 1965 when discussing how users could move between chunks of text, saying this vastly improved the interconnectedness between data and allowed users to build their own experience. He created project Xanadu, a prototype system that stored data, which then allows users to create their own collections using the shared materials (newly added or already in the system) (Nelson, 1987). The description Nelson gave of project Xanadu describes Wikis, especially Wikipedia.

The rapid expansion of the Internet in the 1990s led to the introduction of new technology. For example, the number of users went from 16 million in 1995 to 36 million in 1996; the Internet was becoming the go to place for information (entertainment, business, research, commerce, etc.) (Cohen-Almagor, 2011). In the early 1990s there were a variety of options for finding information (either through searching and/or indexes), these included Internet service providers (like AOL), specific search sites (like Yahoo!), and browser providers (like Microsoft or Netscape) (Green, 2000). However, as the Internet grew, more information was available online causing new problems. For example, the Yahoo! hierarchy directory could not list all the webpages that existed, so it could not always help users find what they needed. The search system, Google, was introduced on the Stanford University website in 1996, it was created by Larry Page and Sergey Brin (both graduate students at that institution) (Cohen-Almagor, 2011). The search algorithm used by Google is still not entirely known, but it uses PageRank, which analyzes the links to and from websites (by how many websites link to a site and the quality of those links) (Green, 2000; Page et al., 1998; Williams, 2017). The information would continue to grow at rapid rates, but technology was slowly catching up or at least providing users with a way to find what they were looking for. Museums would take advantage of this technological evolution and move on to the web.

Digital Museums

It was Andre Malraux who first introduced the concept of a digital museum by describing a "museum without walls" (Malraux, 1965/1967, 12). He discussed how museums distorted the art they displayed, with the focus shifting from the subject matter of the artwork to the artists that created it (Malraux, 1965/1967). For example, a work by Rembrandt or Gainsborough was now more important than the subject they painted or a bowl of fruit could be considered forgettable until you found out Caravaggio painted it. The function of art in museums was stripped away, now it was just a work of art, changing how audiences and artists viewed it (Malraux, 1965/1967). Andre Malraux (1965/1967) said that because of this shift and the introduction of photograph reproductions that "A museum without walls has been opened to us, and it will carry infinitely farther that limited revelation of the world of art which real museums offer us within their walls…" (12). A museum without walls allows users to see and be influenced by works that they have not seen in person, through photography.

It would be roughly twenty-five years before this vision would literally be realized; when museums began using technology like CD-ROM to let users experience the museum at home. For example, in the early 1990s Apple launched the "Computer's Virtual Museum" CD-ROM, with physical museums like the Musee d'Orsay and Louvre following shortly after (Sviličić, 2010). As technology advanced (see above) and the World Wide Web became more popular, digital museums began to appear online. The first digital museums combined static text with images, but as technology developed they began to introduce interactive media (like images users can zoom into) (Styliani et al., 2009).

Examples of early digital museums/exhibitions:

- The Museum of the History of Science (part of Oxford University) created one of the first digital exhibits, "The Measurers: a Flemish Image of Mathematics in the sixteenth century" in 1995. The Museum of History and Science evolved from the Ashmolean Museum (see above) (Bowen, 1999; Bowen, 2010). This site is archived at - http://www.mhs.ox.ac.uk/measurer/text/title.htm
- In 1999 there was the "Remembering Nagasaki" digital museum, created by the Exploratorium (a science museum in San Francisco). This was about the atomic bomb dropped on that city. This digital museum had a section where visitors could leave their own stories/experiences related to what happened that day (Carreras & Mancini, 2014). This website is still accessible (though archived) at - http://www.exploratorium.edu/nagasaki/

- In 1998 the Museum of Modern Art (MoMA) won a Best of the Web awards for their online digital museum, it was simple and artistic much like the museum aesthetic (Bowen, 2000).
- The French Ministry of Culture began an elaborate digitization project in the mid 1990s; their goal was to provide online access to everyone in order to protect objects that were in poor physical condition (Mannoni, 1996; Mannoni, 1997).

Many of the first digital museums were associated with Universities/physical museums because they had easier access to technology (servers, etc.). The number of actual digital museums is currently unknown, but due to the developments in and the decreasing costs of technology, that number is in the thousands, and that estimate has only gone up since 2009 (Styliani et al., 2009).

The World Wide Web is a great tool because it is user friendly, does not cost very much, and yet it provides museum curators with a number of customization opportunities (e.g., designing virtual exhibits with content that is not on display or adding much more information about the content than they could in the museum – the title cards can only be so big) (Styliani et al., 2009). As digital objects increased in popularity, some art historians embraced digital images, others criticized digitization mainly due to slow technology and poor image resolution – this was the late 90s (Cohen et al., 1997). The argument over digitization and the creation of digital museums existed throughout the 90s and early 2000s. However, the argument that digitization would improve collaboration, research, and teaching practices would later be proven true and built upon by experts and novices interested in Art History (Cohen et al., 1997).

Digital Museum Advantages and Disadvantages

Digital museums have both advantages and disadvantages. This includes letting users access the digital museums collection anywhere in the world or the issues that come with digitization and preservation. These websites need to be many things to many people. For example, if a digital museum is associated with a physical museum they need to have the museum's information (hours, location, etc.). If they are online collections they need to follow digital copyright laws or they risk getting the site shut down. If they are used for educational purposes the information that they link to needs to informative, interesting and above all correct. These are just a few of the issues that digital museums face, see below for many more.

Digital museums have advantages and disadvantages, which are commonly related to the audience and physical museums. Here is a summarized list of the advantages found in digital museums:

- One of the main advantages of digital museums is that they can connect objects with supplementary information like images, references, and metadata, which lets users learn about the information that interests them. This enhances learning for both teachers and students.
- They provide users with a wider access to content. They can access it from the comfort of their own home anytime and stay as long as they would like. Most digital museums also provide access to content for people with special needs (those with hearing, visual, and learning disabilities). They do this by working with specialized software and formatting the website, so that users can change it to suit their needs.
- Users can interact with the digital content (zooming in, zooming out, reading additional information about it, etc.), which they can generally not do in the museum.
- They let users move through the exhibit in any way they would like (orderly or randomly for example). Though navigating randomly through an exhibit may make users miss information.
- They allow experts to preserve records online, so even if something deteriorates beyond preservation in the real world there is still a record of it (e.g., the crumbling fresco of Leonardo Da Vinci's *Last Supper*).
- Digital museums can digitize their entire collection for users to view online, providing access to objects not on display.
- Online exhibits could decrease the possibility of damage or theft of the real world content.
- Digital museums can present the object in the environment that they originated from (providing users with contextual information about the objects cultural history). For example, viewing a mosaic in a virtual Roman bath.

And here are the disadvantages found in digital museums.

- Digital museums can mimic a real world experience, but not truly provide one. For example, users cannot touch or feel the objects (though this disadvantage only applies to museums that allow users to touch the artefacts).
- Users do not remember an online visit as well as an in person experience.
- The technology they use can be out-dated or poorly used. Especially the quality of the images – the higher the resolution the longer they take to load and the more likely others will use them (copyright issues), but using lower quality images you risk annoying your audience.
- Each user has different Internet connections and technological experience, so how do digital museums account for all those variables.
- Copyright infringement (especially important in content digital museums that are not affiliated with a physical location).
- The authenticity of the digital museum needs to be evaluated; the information on the sites might be inaccurate.

(Biedermann, 2017; Bowen, 2000; Cody, 1997; Khoon & Ramaiah, 2008; Styliani et al., 2009; Wang et al., 2014; Zhou et al., 2012).

Physical museums were concerned that digital museums could replace or impact visitors going to museums, but this was unfounded. The literature stated that online sites created interest in the physical museum because they provide access to stored objects from the museums collection and they encourage the digital museum audience to visit the real world site (Biedermann, 2017; Karp, 2004a; Marty, 2008; Schweibenz, 2004; Styliani et al., 2009). Digital museums act as supplementary information sources, and unless it is a site with born digital content, they increase interest in the physical objects.

Digital museums use reproduction images (unless the content is born digital) that do not have the same qualities that the physical objects have, Walter Benjamin (1936/2008) called this the objects "aura" (21). The reproduction of art increased once photography became accessible, no longer would art need to be hand drawn or etched in order to reproduce it (Benjamin, 1936/2008). However, photographs cannot replicate the history of the artwork, both the provenance and current physical condition. Benjamin (1936/2008) described this as the "here and now" and it is why technological reproductions cannot represent the authenticity of the object or capture the aura of original work (21). However, technological reproductions can provide users with more detail when viewing the object (zooming in, 3D, etc.) and they can place the reproduction in places that the original cannot go (for example comparing two paintings by Rembrandt that are in different museums in different countries) (Benjamin, 1936/2008; Cody, 1997). Simon J Knell (2003) supports the findings presented by Walter Benjamin, that there is no surrogate for the real object and digital objects (with contextual information) are best used to improve the audiences experience for the physical museum. This also affects digital preservation because then they cannot preserve the essence or aura of cultural heritages objects online (Biedermann, 2017).

Museums collect cultural heritage to control access and preserve objects, but digitization has increased user consumption of information and made reproductions easier to create, transmit, and are of better quality than ever before (affecting the financial gain that these objects represent for museums) (Bertacchini & Morando, 2013; Conway, 2010). Digitization can be both good and bad for museums, good in that when they sell reproduction images they have low transaction and production costs as well as a larger audience who may want to purchase them (Bertacchini & Morando, 2013; Conway, 2010). However, online images can also be distributed without the museums permission, resulting in a loss of income (though watermarks can be added to the images to try and prevent this) (Bertacchini & Morando, 2013). Copyright infringement is an important aspect of digital museums needs to be address.

There are other dilemmas related to digital preservation, not just the cost. Experts need to consider how to combine digitization quality (file format, resolution, metadata, etc.) and preservation quality (Conaway, 2010). For example, how would you preserve audio-visual materials that are currently stored in old technology formats (tape film reels, etc.), do you digitize the media or do you preserve it as it is currently stored? One would think that digitizing would be the best solution, but older technology formats are delicate and can fall apart during this process (Conaway, 2010; Karp, 2004a). Another dilemma is the lack of trained digital preservation specialists, they need to either train new professions (which takes time and money) or outsource the work (which can compromise the preservation of the objects) (Conway, 2010). Whatever preservation specialists are trained, they need to preserve both real world objects and born digital objects to truly capture the current history of cultural heritage online (Conway, 2010).

Digital Museum Audience

Who visits digital museums and their motivation is an important factor for the design of digital museums. The understanding of who uses and why they use digital museums can be used to improve the information architecture of the site (Falk, 2016). For example, by understanding that students are visiting a digital museum to find information and images of artwork, digital museums can present their collection with multiple images or panoramic viewing technology with contextual links for further academic information. What makes the audience even more important in digital museums is that they are not just passive observers, the digital technology lets them interact with the content (Styliani et al., 2009). For example, when a user enters a digital museum they are presented with actions like “Shop,” “Visit,” Research,” etc., which are all actions that they can do with the content of the site. Visitors know that an online visit is different than going to a physical museum, and their expectations change. Because of that, they end up (in some cases) using the site as a resource in their daily lives (Marty, 2008).

Digital museums use social media to interact with their users. These tools can be used for both marketing promotions (tweet how fun you visit was, for example) and as a way to connect with users. It is important to note that social media and the use of technology are creating a sense of expectation for the visitor; they expect the site to be up to a certain standard (both design and technology wise) and for social media features to be available (Nielsen, 2015). That is why many digital museums test out new technologies that promote interaction methods, which audiences can use to visit and experience online content. For example, prototype software was introduced at the Technology and Science Museum “Leonardo da Vinci” in Milan, Italy that allowed users to visit the virtual museum together as avatars in a 3D exhibit (Paolini et al., 2000). While this failed to become a widespread tool, sharing information and objects between the audience and their friends is still important, which is why social media links (to post/tweet/comment on the objects) are found in almost every digital museum.

Visitors' vary between demographics, location, and education (amongst other things). Trying to figure out who is visiting the site is difficult, especially because it can be accessed globally, though websites do have some tools available to them. They can use online surveys and questionaries' to discover why people are visiting the website, or analytics to find out where they go and for how long, or focus groups and interviews to find out qualitative data (Cody, 1997; Falk, 2016; Fantoni et al., 2012). And it is not just one type of visitor that websites need to consider; they need to design the site for as many demographics as possible, as well as for the different types of equipment and internet connections that they may have (Mac vs. PC or rural vs. city connection speeds) (Cody, 1997). The subject of user equipment/technology also brings up different technological skills users have (novice, expert, or somewhere in between), digital museums need to design their sites so that they don't overwhelm the novice user, but also don't annoy the advanced (Cody, 1997; Walsh et al., 2016). There is much to consider when designing a digital museum for their target audience.

What digital museums should do is try to narrow down their target audience using user studies (some of which were mentioned above) in order to determine how to tailor the design of their site for them (Bowen et al., 1998). For example, brochure museums should make sure that the physical museums hours of operation, location (with directions is helpful), and contact information is easily accessible. Learning digital museums should define exactly whom they are looking to educate (and it could be more than one type of user) and provide the resources they may like (e.g., children might like educational games about the content they are looking at). And content digital museums should make sure they have the most important works of art from the period or theme that they are representing (Barton, 2005; McDonald, 2005; Piacente, 1996; Schweibenz, 2004a, Styliani et al., 2009; Teather, 1998). These are just a few of the ways that museums can design their site for users. A further consideration is defining the types of users and their motivation(s) for why they use digital museums.

There have been studies that try to define the different digital museum users; the findings of these studies are presented differently with different terminology and/or broad vs. specific categories. For example, Styliani et al. (2009) said that digital museums have three different audiences – researchers, students, and tourists. When looking at digital museums you can see how they design different sections of the site for those three types of users. However, these could be further clarified, for example what type of "student" are they talking about (elementary, high school, college, etc.) and what about teachers as a visitor group (or are they under researchers). The reason teachers are specifically mentioned is because many websites (including a few in this study) have a dedicated area for this group.

Another study conducted a literature review in order to identify the main user groups of digital museums. David Walsh, Paul Clough, and Jonathan Foster (2016) found that users could be grouped under generalized categories organized by their knowledge of the subject matter within a digital museum's collection:

- Expert or professional users (these are trained professionals like curators, archaeologists, and historians).
- Hobbyists or semi-experts (they have a keen interest in the subject matter and some familiarity with it but they have not been formally trained and are there to learn more about it).
- Non-Experts or Lay users (people who have interest in the subject matter but they do not know that much about it yet).

(Walsh et al., 2016, "Broad Categories of User").

These users groups could also be defined by their motivation for visiting the digital museums. In the above instance they are defined by their learning motivation. Walsh et al. (2016) additionally describe different user groups based on the actions they are most likely to perform - general visitors (people looking for the hours and direction to the physical museum), educational visitors (those looking to learn something new), and specialist visitors (researchers who are looking for detailed information about the collection) ("Other Groups of Users"). These closely relate back to Styliani et al. (2009) conclusion that there are three types of users (researchers, students, and tourists).

User studies have been conducted to determine the motivations of users visiting digital museums. For example, the Indianapolis Museum of Art digital museum used a survey questionnaire to determine why users were on their site (e.g., were they there to buy something or just casual browsing information on the site) (Fantoni et al., 2012). The majority of the respondents (50%) said that they were there to plan a visit to the museums, followed by 21% saying that they were there to find specific information that they were interested in (Fantoni et al., 2012). The motivation and types of users interconnected (the "tourist" user's motivation is in the definition).

There are also visitors who are motivated by the tools and technology that digital museums provide for their users. Three types of visitors put forth (from a survey of 64 museums – including art, science and history museums) by Mathilde Pulh and Remi Mencarelli (2015) are "communicator-visitor," "curator-visitor," and "artist-visitor" (44). These types of museum visitor are defined by what they want to do when on the site. For example, the artist-visitor wants to create and interact with artwork on the site; the curator-visitor wants to create their own collections of object that interest them; and the communicator-visitor uses social media (including the websites blog if applicable) to communicate with the museum and about the collection (for example posting on the websites Facebook page) (Pulh & Mencarelli, 2015).

For more information about the audience of the digital museums evaluated in this study see Chapter 4. They were defined as part of the contextual framework for each digital museum, identifying the users, content, and context to see how the information architecture should serve the website (for more information about contextual framework see below).

Information Architecture

Information architecture (IA) primarily organizes and structures the information of websites. Users may never notice its existence, but they use IA features every time they are on the Internet or in a digital environment (like a word processor). You can also find information architecture in the real world, for example how a grocery store organizes, labels, and structures shelving units so that users can easily find things (Covert, 2014; Rosenfeld et al., 2015; Spencer, 2011). This study will focus on the digital side of IA. The definitions of information architecture by different authors will be outlined as well as a closer look at why this field is important, the main systems that belong in IA, information architects, and how they conduct IA research and design. The contextual framework (the users, content, and context specific to a website) will be outlined, providing details about why this is so important in the design of IA. For additional background information, a brief history of IA will be detailed, based on how information architects influenced the development of the field. This section will conclude with a look at how standardizing information architecture components have been dealt with in previous studies.

Definition

The definition of information architecture differs depending on the author, but most of the sources examined use the definition (or a variant of) provided by Rosenfeld et al. from *Information Architecture: For the Web and Beyond* (2015). This definition is separated into four points because IA is a complex field that cannot be easily defined. Louis Rosenfeld, Peter Morville, and Jorge Arango's (2015) definition will be used for this study, and it is:

- The structural design of shared information environments
- The synthesis of organization, labeling, search, and navigation systems within digital, physical, and cross-channel ecosystems
- The art and science of shaping information products and experiences to support usability, findability, and understanding
- An emerging discipline and community of practice focused on bringing principles of design and architecture to the digital landscape

(Rosenfeld et al., 2015, 24)

This definition describes how information architecture works in digital environments by structuring and/or organizing information across multiple platforms, so that users can find and understand the information presented online (Rosenfeld et al., 2015). The importance of making websites work for users (across platforms) is a founding principle of this field; this will be covered in more detail under contextual framework below.

Abby Covert (2014) describes information architecture as a concept that "… make[s] sense of messes caused by misinformation, disinformation, not enough, or too much information" (3). This is especially important considering how much information is on the Internet and in other digital environments. Covert (2014) goes on to say how information architecture organizes the different parts of something into an understandable unit, which includes determining how to structure the information so that the website's message can be communicated to users. For example, organizing labels in the global navigation bar links pages together and describes the organization scheme of the site. Donna Spencer (2011) provides a similar definition, in that information architecture organizes information, labels that information, and provides ways to access that information in whatever ways that make the most sense for the target audience and the content in the site (Spencer, 2011). This is generally for websites and intranets, but IA principles can be applied to the real world (e.g. a grocery store as mentioned above). Both definitions discuss the organization and access of information; this is a common theme in all the definitions.

Wei Ding, Xia Lin, and Michael Zarro (2017) define information architecture as the organizing and structuring of information to create clear and understandable interfaces and systems for users so that they can find, learn, and manage information in a website (Ding et al., 2017). Information architecture also deals with multiple platforms in order to allow for easy access to information and provide consistency to users (Ding et al., 2017). Much like the definition given by Donna Spencer, Ding et al. (2017) describe how IA serves the needs of the users by organizing and providing access to content. Both of those descriptions are similar to the four-point definition given by Rosenfeld et al. (2015).

There are many more definitions of information architecture (from numerous sources) that contain the same themes – the organization and structure of information to provide access to content for users:

- Information architecture is a model that describes, creates, and maintains how websites organize, contextualize, provide access, and present information for users (Crawford, 2011).

- "Information Architecture (IA) is the art and science of structuring and organizing information environments to help people effectively fulfill their information needs." (Toub, 2000, 2).
- Information architecture is a craft or an applied art that is focused on the organization and access to information (Resmini & Rosati, 2012). It relies on collaboration, supporting literature, guidelines, and best practices to create and advance the field (Resmini & Rosati, 2012).
- "Ideally, it is the information architecture that defines the service, clarifies the vision, determines content and functionality, specifies how users will find information, and maps how the services provided will change and grow over time, all dependent upon the needs of users." (Simon, 2008, "Introduction")
- Information architecture organizes, structures and labels content in websites to help users find the information they are looking for and/or to complete tasks (for example buying something) (Usability.gov, 2013).

You can see the similarities between these IA definitions and the one put forth by Rosenfeld et al.. Though Rosenfeld et al. (2015) does not specifically mention the term "user" they do discuss designing websites to support "… usability, findability, and understanding" (24), which are directly related to how users experience a webpage. For example, Stephanie Crawford (2011) discusses how websites present information for users and Andrea Resmini & Luca Rosati (2012) discuss access to information (even though, like Rosenfeld et al., they do not explicitly say the term user, but that is who would access the information). There are a few reasons why these definitions are similar. Rosenfeld et al. are well-respected experts in the field of information architecture and have been for many years. Their book (the polar bear book) has published four editions and is considered the IA bible (The Institute of Information Architecture, 2017, What is Information Architecture?), so it makes sense that other IA scholars adopt their definition. This also could be because information architecture has a general consensus between scholars about what they do and why they do it (for more information see below). For example, the organization of information is mentioned by almost all of the definitions.

The Importance of Information Architecture

Why is information architecture important? We began to answer this question while discussing IA definitions (for the organization of information and so that users can find, understand and interact with that information). IA solves many different problems that are found in digital environments.

Louis Rosenfeld, Peter Morville, and Jorge Arango (2015) said that information architecture tries to fix two issues in the digital world – information overload and how to present information across devices/platforms for consistency. IA tries to solve information overload by organizing and structuring information so that it does not overwhelm users (Rosenfeld et al., 2015). For example, providing global and local navigation that is consistently placed and accessed, so that users always know how to travel to different webpages (or back to the homepage). IA works to solve issues of information access between the different devices/platforms (like a smartphone not being able to access the menu navigation because the site does not scale for the small screen) by providing consistency in structure and design across these platforms (Rosenfeld et al., 2015). Though in order to do this you need to understand how each platform relates to one another. Solving these two issues is so important; Rosenfeld et al. included it in their IA definition (points 3 and 4). Abby Covert (2014) agrees with Rosenfeld et al., she said information architecture solves the issues that cause confusion in websites – if there is too much information or not enough, if the information is confusing or not what the user is looking, and/or a combination of those (8). There is more accessible information than ever before (this influx of information was covered in the technology section under technology) and IA must make sure that they structure and organize information for users in the clearest way possible.

According to Christina Wodtke and Austin Govella (2009) information architecture improves three characteristics in websites – "the findability, usability and understandability" (also discussed in the Rosenfeld et al., 2015 definition above) (282). These three characteristics or components of information architecture design focus on the website's audience/user to make sure that they can find what they are looking for, accomplish what they want to do with that information, and understand the information they find (Wodtke & Govella, 2009). That is why

IA is so important to websites, especially sites like digital museums that exist to store and provide information to users – access and understanding is key.

Donna Spencer (2011) describes two reasons why IA is so important, it determines how easy a website is to use, and websites with a well-designed IA give users confidence in what they are doing (when they can find and understand the information easily). This is reasonable, if a website is clear, learnable, and understandable you gain confidence in both yourself and the website. This would then motivate users to return to that website and probably explore the different information it contains.

If IA can make users think positively about a site, it can also make them think negatively about it. If websites are disorganized, badly labeled, or just generally display poor IA design, the user is probably not going to have a great experience using that website and unless it is a necessary portal in their everyday life (like intranets) they will not be back. For example, "A disorganized intranet eats up employee time while employees try to find information they need, a disorganized external Web site prevents online shoppers from finding the items they want to purchase…" (Wodtke & Govella, 2009, 281). Steve Toub (2000) also discussed the impact poor IA design has on both the user and the business it represents. For example, failed searches or complex forms can cause a loss in revenues for a business because users cannot find what they want to buy or are too impatient to create a user profile before buying something (Toub, 2000). Sometimes people do not want to sign up to another website in order to buy something, especially if it is the first time they are using that site (you need to test out service, delivery, etc. before committing to an ecommerce site). Intranets also have issues with organization, but instead of a loss of revenue (although that is by-product of not having staff work to their full potential) it is a loss of productivity (Toub, 2000). No matter what the problem is with the information architecture, if there is an issue, it needs to be fixed to produce the best results for a website.

Information Architecture Systems

The information environment users encounter when searching/browsing, impacts how they understand something (through design, structure, organization, labeling, navigation, etc.) (Rosenfeld et al., 2015). There are five main components that make up websites information architecture. These are organization, navigation, labeling, search, and vocabulary systems. These systems dictate how users find, understand, and use the information in websites, just like in the definitions discussed above. A brief overview of these systems will be provided below, for more information see Chapter 4.

Organization systems structure and classify information in a website. There are two main components, organization structures (which are made up of different possible structures, e.g., hierarchy or database oriented) and organization schemes (which explain the different ways information can be classified, e.g., alphabetically or by task) (Covert, 2014; Crawford, 2011; Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). When organizing content in a website or other digital environment, users should be able to tell where they are, and what information is available in the websites (Wodtke & Govella, 2009). There is so much information on the web, the design of this IA system is becoming more and more important.

Labeling systems describe chunks of text that are arranged according to the organization system. There are two types of labels – textual and iconic – which also generally represent navigation links (Rosenfeld et al., 2015). The labels for a website should be simple, represent the websites target audience (but still use correct terminology), and consistent across the site (Crawford, 2011; Rosenfeld et al., 2015; Spencer, 2011). For example, calling the objects in digital museum a "collection" in one area and "artwork" in another area is a consistency issue, which can cause confusion (even if it is only momentarily) that distracts the user.

The navigation system is how users move through a website to reach the information/content it holds. They do this by clicking on links found throughout websites, including global, local, and contextual navigation (Crawford, 2011; Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). When designing a navigation system in a website you should consider how the content is organized, where users need to go, and what should they be able to do (Wodtke & Govella, 2009). For example, contextual links represent connections between objects and contextual information; these are added to objects in the site in order to get users to click on them (Cody, 1997).

The search system is an important tool in websites. Most users expect to find one in a website, though this does not mean that you need a search system (make sure it works for your site) (Rosenfeld et al., 2015; Spencer, 2011). The search system should seem simple to the user, just type something, press enter, and get the results. However, in reality there is a lot going on in the background (algorithms, content indexed for searching, how results are displayed) (Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). For example, when you search for "Paintings" the algorithm of the search system queries the metadata and/or full-text of the content (this depends on what has been indexed for searching) to return the results in a grid pattern with select information displayed.

The vocabulary system allows "…you to structure and map language so that people can more easily find information." (Rosenfeld et al., 2015, 309). Examples of the vocabulary system include controlled vocabularies, thesaurus, and metadata (Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). For example, the controlled vocabulary can be seen in the consistent way a site structures artist names (e.g., is it last name then first name or vice versa).

These five systems constantly overlap. For example, labeling plays a major role in defining the organization schemes of a site, telling users what navigation links are available, and they are generally made up of terms from the vocabulary system the site uses. This overlap can make planning, designing, and implementing information architecture difficult, but it is also the reason why it is so important to get it right (one bad component can ruin the rest of the site).

Information Architects

Information architects are the people who do the planning, designing, and implementation of the websites IA. They define how the website will work, how users will find and understand the information on the site (using the systems discussed above), and how the website will balance the needs of the user, content, and business goals (Rosenfeld et al., 2015, Simon, 2008; Spencer, 2011). Richard Saul Wurman (1997) defined information architects as those who organize patterns in data and structure so users can find what they are looking for. This man is one of the founding fathers of information architecture (see more below). Information architects are also responsible for building bridges between different components in the IA design process, including between different platforms that websites are available on, between the users and the content (how they access it), and between researching and implementing the IA for a site (Morville, 2011). According to Donna Spencer (2011) many different professional groups can design the information architecture of a website – usability specialists, web developers, project managers and anybody who is interested as long as they take the time to learn about it (though having communication, language, and attention to detail skills would help)(Spencer, 2011). While this is certainly true (this author does not have any formal training in the field) being informed about information architecture and how to practice it through research and design is important.

Information Architecture Research and Design

Information architecture does not just magically appear in a website, it evolves from hours of research and design. While researching the information architecture of websites, it is important to define the contextual framework (users, content and context) to get to know the details of the site (this will be discussed in more detail below). The research will be summarized into a strategy that outlines how the information architecture design will be implemented in a client's site (this includes a report, wireframes, etc.) (Ding et al., 2017; Rosenfeld et al., 2015). There are many different types of research that information architects can conduct to discover the IA components a website needs.

When an information architect begins a project, it is a good idea to gather up any pre-existing research available to them. This includes items like existing content, inherited technology, business mission and goals, and the current IA design (Crawford, 2011; Spencer, 2011; Toub, 2000). Once they have reviewed that information they can begin to conduct new studies in order to see what works (and what does not) for a site. For example, they can conduct a heuristic evaluation of competitor's sites to see what IA features they use and then either choose to use a similar structure or use this evaluation to view what does not work in a competitors site (this type of evaluation can be conducted at any stage of the website IA design) (Toub, 2000; Rosenfeld et al., 2015).

The research stages should be used to identify the target audience for the website. This can be accomplished using both qualitative and quantitative methods like interviews, focus groups, online surveys, web analytic software (if there is a live website), and much more (Ding et al., 2017; Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). Once user research is completed, analysis of the data can begin, this can involve creating personas that represent the typical user of the site and scenarios of how they will use the site (Rosenfeld et al., 2015; Spencer, 2011). For example, if "tourist" was determined to be a typical user for a digital museum, you would build a profile of a typical tourist user that would include a name, image, and background information. Then you could have that user complete a task on the website, such as finding out when the museum opens on a Thursday.

Once initial research is completed you can begin analyzing the data and documenting it to present to the clients (if you have them). This can include things like creating diagrams to show the organization structure and navigation system that will be created for the website. The diagrams that can be used include flow diagrams, block diagrams, Gantt charts, Venn diagrams, hierarchy diagrams, and wireframes (Covert, 2014; Crawford, 2011; Ding et al., 2017; Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). This stage will need to have several evaluations to make sure that the diagram works for the users, content, and business (Spencer, 2011). When the design stage is completed, the results are presented to the client (if applicable) and a prototype created (digital or paper) (Ding et al., 2017; Rosenfeld et al., 2015; Spencer, 2011). This will be tested again in an iterative process until the live website is launched (and even then there will be more testing).

Contextual Framework

The definitions presented above describe users as the foundation for information architecture design, but there are two other important components – the context and content of websites. The contextual framework is made up of those three groups. These three factors drive an information architecture design project and because each one depends on the other and they generally vary between sites (Rosenfeld et al., 2015). These three components are defined as (summarized from multiple sources),

- Users are the people who use the website or digital environment that the IA is being designed for. Consider the type of audience, their motivation, level of technical expertise, and their experience with the subject matter when planning out the IA of the site. You also need to figure out what they will be looking for and what they want to do when they are on the website.
- Content is the information users are looking for on the website. This can be images, documents, items to purchase, and much more. It is not just types of content that needs to be considered, but also how the site presents that information to the user. Are they in a list, grid, index, etc. and what information is displayed about the content (metadata)?
- Context refers to the business side of the website. You need to know the organizational goals, business model, mission, technology constraints, budget, etc. and how these effect the website IA design. The stakeholders will have opinions about the design of the site (and they should be considered, right or wrong, in the design phase).

(Covert, 2014; Ding et al., 2017; Resmini & Rosati, 2012; Rosenfeld et al., 2015; Simon, 2008; Spencer, 2011; Toub, 2000; Usability.gov, 2013; Wodtke & Govella, 2009)

You need to know a lot about these three groups for the IA to work, which should be your first priority when researching the IA needs of a website. When evaluating the nine select digital museums, a contextual framework was created for each site in order to become more familiar with the websites. The user of the site was defined by the content and how it was presented, as well as if there were any audience specific organization schemes. The content was explored and then listed (e.g., artwork, blog articles, contextual documents, etc.). The context of the website was determined by looking at the mission statement and any other organizational information available on the site.

History

The history of information architecture truly began with the invention of modern technology, the computer and Internet (for an overview of this history see above). However, Richard Saul Wurman first used the term "Information architecture" in 1975, in an article co-written by Joel Katz called *Beyond Graphics: The Architecture of Information* (Resmini & Rosati, 2012; Simon, 2008; Wurman, 1997). In 1976 he would use this term during a presentation at the American Institute of Architects, popularizing it (Crawford, 2011; Resmini & Rosati, 2012; Simon, 2008). When Wurman described information architecture, it was not how we think of it today. His version of information architecture was more closely associated with visual design, not the structure of a site (Morville, 2004; Resmini & Rosati, 2012; Wurman, 1997).

While the creation of the term "information architecture" is credited to Wurman, there is some debate about that. In 1970 Xerox Palo Alto Research Center (PARC) wrote a charter to develop technology in order to support "architecture of information" (Pake, 1985 as sited in Resmini & Rosati, 2012). Some information architects say that this is when and where the term truly originated. Nonetheless, the field of information architecture was established and would continue to grow. In the 1980s information architecture was not that popular or well used, in this era it was more closely associated with computer infrastructures than with website design (Resmini & Rosati, 2012). The information design from the 60s and 70s combined with the computer systems design (the 80s) merged into what we now know of as information architecture (in the 90s) (Resmini & Rosati, 2012). Two of the most influential information architects gained notice in 1993 with the founding of the IA firm Argus Associates – Louis Rosenfeld and Peter Morville (Morville, 2004; Resmini & Rosati, 2012). These two men changed the practice of information architecture by merging it with Library and Information Studies (LIS) ideas; it was here that the "systems" (organization, labeling, navigation, search, and vocabulary) were introduced to the field (Morville, 2004; Resmini & Rosati, 2012). And they would popularize IA with the publication of their book *Information Architecture for the World Wide Web* (the first edition was printed in 1998), which is also called the polar bear book (Morville, 2004; Resmini & Rosati, 2012). The history of IA was not without its challenges, the dot com bubble burst in 2001 saw the closing of Argus Associates, budget cuts, and people fired

(Morville, 2004). However because IA is so closely related to the usability of a website, advancements in technology meant that there will always be a need for it to continue on.

As the amount of information on the Internet increased, the need to organize and access it became very important. In some cases this resulted in new technology like Google to provide information retrieval, in others it required the organization and structure of the information in websites. It is the latter that popularized IA. As technology advanced, there was a shift from focusing on information architecture just for computers to include new platforms like smartphones, tablets, watches, etc. (Ding et al., 2017; Resmini & Rosati, 2012). Information architecture was now a powerful field of study that integrates multiple platforms with a consistent organization structure across them, while considering the linked data between them (Ding et al., 2017; Crawford, 2011; Rosenfeld et al., 2015). This also meant that more attention is now being paid to social objects like tagging and social media sites (Morville, 2011). Users are on the web constantly (to the point that it defines our lives), information architecture does and needs to play a large role in this new reality (Resmini & Rosati, 2012).

Standardizing Information Architecture in Digital Museums

Information architecture (thesauri, schemas, metadata, etc.) needs to be a priority when designing a digital museum, particularly for users to access and understand the objects presented in the site (Teather, 1998). One way to make information architecture a priority is standardizing the different IA systems (organization, navigation, labeling, search and vocabulary systems). Standards are very important when implementing an information architecture design (or any website component), they can promote access to information for users and enables interoperability (Riley-Huff, 2012; Simon, 2008). Standardizing information architecture lets information architects make changes to the different systems without affecting the whole website (Simon, 2008). For example, when they know and are familiar with the IA components (those that have been standardized) of a site they can easily go in and make changes (deleting, replacing, adding new items, updating, etc.) (Simon, 2008). Though standardization is important, the literature often only focuses on metadata and controlled vocabularies (vocabulary systems) when discussing it in digital museums. This study will create criteria for IA design in digital museums, the first step in creating IA standards.

Standardization of metadata and controlled vocabularies is important to ensure interoperability between cultural heritage websites; this will help provide universal access to knowledge (Srinivasan et al., 2009). Although there is not a universal standard yet for digital museums, the most common metadata standard is currently Dublin Core (DC) (Zhou et al., 2012, 220). However, Murtha Baca (2003) states that Dublin Core is not the best option for digital heritage objects as does not cover the full scope of what they need in their metadata (etc.). This is why standards are so difficult and necessary, there is a lack of agreement between experts in the field about what the standard should be, meaning that interoperability will not be achieved until there is a universally accepted standard. However, standardization between different metadata schemas can be achieved with metadata mapping, this compares metadata elements from different metadata schemas in order to isolate the similar fields and connect them when a search occurs across different databases (Baca, 2003). This is a difficult and complex process that needs experts to create the maps, and again the experts do not always agree with the choices made during this process.

Pre-existing controlled vocabularies are another type of standardization, using predetermined complex terms that contain the same structure across sites. The best standardized controlled vocabularies to use when inputting metadata are those that have been created especially for cultural heritage objects like the Art & Architecture Thesaurus by the Getty Research Institute (Baca, 2004). However, this is not always possible if they are working with pre-existing metadata schemas/controlled vocabularies or if the organization that the site is connected to (either a physical museum or a digital company) uses a different schema.

Whatever metadata schemas and/or controlled vocabularies are used, standardization of metadata schemas has both advantages and challenges associated with it. They are:

- Search systems can be used for information retrieval across different databases.
- Cultural heritages organizations can share data.

- Users can navigate and search better on websites if they are familiar with the metadata and controlled vocabularies (and the more sites that standardize their controlled vocabulary the easier it will be to learn).
- Experts from different cultural heritage organization (libraries, museums, and archives) can learn how other people utilize the standard.
- The users do not get a say in the terminology, just the experts who create them.
- The standardization can pick up the biases of the people who create the metadata and controlled vocabulary for the site.

(Srinivasan et al., 2009)

Standardization is a good idea for many reasons. Creating the criteria is the first step to create standards, which are necessary to provide interoperability between cultural heritage sites. This will make organizing content easier and access to it more consistent for users.

Chapter Summary

Digital museums and information architecture are the main components studied in this thesis. These components have been defined and placed in a historical context, so that the reader can learn more about them and understand their significance to this study.

Digital museums are an online environment that uses different technologies (like 3D graphics and multimedia) to present collections of objects and contextual information to create an experience for users (Foo, 2008; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009; Zhou et al., 2012). There are different types of digital museums (brochure, content, and learning) found on the Internet, but all of them can be classified under that definition. The history of digital museums can be seen as two historical timelines merging. The history of the museum moved from private or upper class accessible collections to open access (digital museums). The history of modern technology saw the invention and evolution of the computer as well as the development of the Internet. The history of digital museums included the utilization of CD-ROMS before taking advantage of the Internet and moving online. A closer look at the advantages, disadvantages and audiences of digital museums was then provided.

Information architecture is a robust field that organizes and provides access for users to digital content. An overview of information architecture was presented to round out the IA definitions discussed. This includes why the field is so important, the different IA system components, and what an information architect does as well as how they do it. The users, content, and context of a website were defined and placed in the broader category of contextual frameworks, which determines how a website's IA should be designed and implemented. Influential information architects, chiefly Richard Saul Wurman, Louis Rosenfeld, and Peter Morville were responsible for the development and evolution of the history of information architecture. As information architecture developed, the idea of standards was introduced. Primarily (in the literature) standardization applies to the metadata and controlled vocabulary of a site. However, all IA components can be standardized, which will improve consistency and access between similar websites. The list of criteria in this study is the first step in creating IA standards; users studies and additional research will be required to create standards.

The information architecture of digital museums will be evaluated for this study. The next chapter outlines how this will be accomplished.

Chapter 3: Methodology

Introduction

The goal of this study is to discover how digital museums implement information architecture in their website and if this information can be condensed into a list of criteria. The list of IA criteria will help improve the structure and labeling in other select digital museums as well as promote consistency and access to information for users. To do this, the study needs to establish what the information architecture principles are and how they are implemented in real life. The research design for this study is divided into two parts. The first is a content analysis study, which will examine general-knowledge information architecture literature to determine the IA principles. The second is a usability inspection method, heuristic evaluation. The heuristic evaluation examines nine digital museums to determine (a) what information architecture principles they use in site organization and structure, and (b) if there are any errors or problems with the implementation of those information architecture principles. By applying these methods to IA literature and live digital museums, this study can theorize what IA principles (the criteria) should be found in select digital museums and offer suggestion for how to fix any IA problems discovered.

Research Questions

The questions that this research study will set out to answer are:

1. Can the current information architecture design in digital museums be condensed into a list of criteria to promote consistency between like websites and ease of use?
2. What are the information architecture principles present in the select information architecture literature?
3. How do select digital museums use information architecture principles in their live website design?
4. From the results of the study, what information architecture criteria should digital museums have?

To determine if digital museums effectively organize and present information to users, two research methods will be used. A content analysis study will determine the information architecture principles found in IA literature. These principles will be defined to create a set of heuristics for the next portion of the study. A heuristic evaluation will use the heuristics previously created in order to evaluate nine select digital museums to uncover how they use those principles (and they may not use all of them). As well if there are any bugs/issues related to the information architecture of the site. When all the digital museums have been evaluated, the digital museum reports containing the information architecture principles discovered will be reviewed, compared, and contrasted between the different types of museums and then between all the select digital museums studied. The secondary comparison will create a list of criteria that select digital museums can implement to provide consistency between similar sites. An analysis of any bugs/issues found in the site will be completed as well, to help improve the existing information architecture. The decision to examine information architecture, instead of graphic design or interaction design, was made because providing access to information is an important characteristic that digital museums and information architecture share.

Research Approach

This study follows the interpretive paradigm approach, meaning there are multiple realities or interpretations for the phenomenon, bound by the context that surrounds what is being studied (Merriam & Tisdell, 2016). Studies using this approach collect data in an inductive process in order to create and answer their theories (Merriam & Tisdell, 2016). The interpretive approach is associated with qualitative studies.

This is a qualitative study. Though both quantitative and qualitative methods research phenomenon that occur in the real world, including technology, and recognize that the phenomenon being studied has many layers and needs to be portrayed in all its complexity and so I chose to do a qualitative study, providing qualitative interpretation to the research (Leedy & Ormrod, 2010). Digital museums are multifaceted structures with information architecture found in every level, so an in-depth examination is necessary. This is not to say that only a qualitative approach (or the selected methodology) would work for this study, but it allows every layer of the phenomenon to be explored in a flexible way.

In the qualitative approach, the researcher is an instrument of the study – they theorize, collect, and interpret the data (Connaway & Radford, 2017; Leedy & Ormrod, 2010; Tracy, 2013). The background and personal views of the author are part of the research, because they impact the research questions and theories. Inductive reasoning combined with personal insights are a resource during this type of study, and as long as this is clearly stated in the research design it is an asset (Connaway & Radford, 2017, 215; Tracy, 2013, 11). Biases in this study are discussed in the validity and reliability section of this chapter.

Content Analysis

Content analysis is the first method used for this study, conducted during the initial literature review. This procedure will analyze information architecture literature in order to discover key principles. Content analysis is a close and thorough examination of documents, enabled by a set of predetermined codes (created prior to the study), which directs the extraction of information according to categories and themes (classified by the coding) (Bengtsson, 2016; Hsieh & Shannon, 2005; Mayring, 2000; Schreier, 2012; Wilson, 2011). This method uses a structured process, which changes depending on the "type" of content analysis performed. Content analysis was selected for this study in order to create a list of heuristics for the heuristic evaluation (stage two of the research study).

A formal content analysis study was used to ensure that the findings are as precise as possible. The other option for this stage of research was a review of the relevant literature (Leedy & Ormrod, 2010). Following this structured method, the decisions made during the study of the information architecture principles were catalogued and presented for review. Content analysis emphasizes the validity and consistency of the method findings; this is necessary because the heuristic evaluation is built on the information architecture principles established during the content analysis. A review of the literature would not have had the same formality and the creation of the heuristics would have been much more affected by personal biases.

Content analysis is divided into two main approaches. These are defined into different categories (given different names by different researchers) –latent/manifest, inductive/deductive, or qualitative/quantitative (Bengtsson, 2016; Hsieh & Shannon, 2005; Potter & Levine-Donnerstein, 1999; Schreier, 2012; Schreier, 2014). The differences between the content analysis strategies are as the names suggest – one is structured (manifest) and one is flexible (latent) in analysis interpretations. The manifest/deductive/quantitative content analysis uses predetermined hypothesis and categories to systematically examine the texts studied (concept driven) (Schreier, 2014; Potter & Levine-Donnerstein, 1999). A latent/inductive/qualitative content analysis is a data-driven approach, meaning the patterns and themes of the text are subjectively extrapolated in order to reach a conclusion (Schreier, 2014; Potter & Levine-Donnerstein, 1999). Hsiu-Fang Hsieh and Sarah E. Shannon (2005) propose a third type of content analysis, a “summative” approach that studies individual words in a text to explore the frequency and usage of them (categorized under the latent method). This study follows the “latent” design because the majority of the categories are derived from data analysis of the pilot studied coded data.

All types of content analysis follow a specific research structure. Summarized from several sources (Bengtsson, 2016; Leedy & Ormrod, 2010; Schreier, 2014; Schreier, 2012; Krippendorff, 1989; Wilson, 2011), the research structure is:

1. Define the research questions and determine the aim of the study
2. Select the sample material
3. Begin building the coding frame
4. Break down the sample into units of analysis
5. Begin a pilot study to test out the coding frame
6. Review and update the coding frame based on the pilot study results
7. Analyze the sample
8. Interpret and present the findings.

Though this is a structured approach to analyzing text, the actual collection of data can be subjective. It depends on the type of researcher and the type of research.

The selected sample is four general knowledge books. They are - *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella; *A Practical Guide to Information Architecture* by Donna Spence; *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango; and *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert. The books selected were the first four choices when the search string "information architecture" was entered into the goodreads database. The newest edition of the four search results was selected. The reason why this search was conducted on the goodreads site is because the literature will then have a broad appeal with the user base. The search string "information architecture" was used so that the samples would represent broad general knowledge information. Books and not articles were chosen because the topic would then be explored in length. This is not a representative sample, because the number of information architecture literature is unknown.

Choosing books from a popular database is not the ideal selection process and introduces subjectivity into the research (as well as limiting the generalizability of the study) (Lacy, Watson, Riffe, & Lovejoy, 2015). To combat this sample selection decision, it was suggested that a precision/recall calculation be completed. There were 51 results returned by the search and 35 of the search results were relevant sources, so the precision of the database is 35/51= 68%. The 16 sources that were not relevant included both items that did not apply to digital environments (such as *Geologics: Geography Information Architecture* – a city planning book) and multiples copies of the same book. The recall was not calculated because a complete list of all the relevant sources on goodreads is unknown. While 68% precision in not a great result, the samples represent general opinion and they had the most ratings by readers (the lowest being 154 user ratings with the average rating of 3.76 or better) in all the results. This is a non-probability sample, which means that the results cannot be generalized to all information architecture literature (Leedy & Ormrod, 2010; Merriam & Tisdell, 2016; Tracy, 2013). This study examines the books by chapter, but the preface, table of contents, appendices, index(s), and reference/bibliography sections will not be included.

Creating the coding frame is the next step in the content analysis method. The coding frame is the structure for the text analysis, the main categories or themes are stated with descriptions (examples, units of analysis, etc.) and coding instructions provide rules for how to (and how not to) code the text (Bengtsson, 2016; Lacy, Watson, Riffe, & Lovejoy, 2015; Leedy & Ormrod, 2010; Mayring, 2000; Schreier, 2014). This is the most important part of the content analysis, it determines if you have an accurate and reliable study. To create the coding frame (or scheme) for this study, the main categories were selected – organization, navigation, search, labeling, and vocabulary systems. These categories (and future subcategories) need to be clearly explained (especially in multi coder studies). This includes a definition, example, and identification code (Bengtsson, 2016; Mayring, 2000; Schreier, 2014). Prior to the pilot study the rules of analysis will be outlined, including a statement of the unit of analysis, and variables of interest (see Appendix B). The variables of interest define what information in the text is not covered, for example, in this study graphic design and discussions about the user will not be coded because it is beyond the scope of the research (Schreier, 2012).

The study guidelines influence how the coding frame is created. For example, you cannot have one passage of text coded into two subcategories under the same main category (they can be coded more than once, but into subcategories organized under a different main category) to remain mutually exclusive (Bengtsson, 2016; Schreier, 2014). The coding frame will support these rules. When coding the text in the pilot study, all categories will be coded with the first letter of the theme (e.g., O, N, S, etc.). In qualitative studies, the pilot study can be done on one of the samples in the project as long as all aspects of the research question are represented, this sample will have to be redone after the coding frame is revaluated (Mayring, 2000). Once the pilot study is finished, these codes will be examined for subcategories in order to create the full coding frame (these will also be defined). After the coding frame is completed, the study will begin with the examination of the complete sample.

Data collection will examine each of the four books selected for the study. During data collection the text will be broken down in order to discover patterns and relationships (though this will result in some loss of context) (Schreier, 2014). To establish information architecture principles for the heuristic evaluation, the literature was broken down into chunks to understand the components found. Every chapter will be analyzed; any text about an information architecture principle will be excerpted and coded in a report form (see Appendix C). All information applicable according to the coding frame will be coded. Once the four information architecture books have been analyzed, the coded data will be edited to make sure that the coding is correct and that the data followed the rules of the coding frame.

Data analysis will begin with the creation of the heuristics. The coded data will be examined and any key excerpts identified (definitions, examples, components, etc.), this will be compiled into a list for the heuristic evaluation. Both qualitative and quantitative methods will be used in the analysis of the data collected. Qualitatively, relationships between these categories will be analyzed to determine how they work together between the sources. Quantitatively, statistics like code frequencies will be used to provide additional insights on the coded relationships. (Leedy & Ormrod, 2010). The findings will describe the information architecture principles discovered and identify any patterns using examples and statistics to present a full picture of the study.

Heuristic Evaluation

The second methodology is the usability inspection method, heuristic evaluation. When evaluating a user interface there are four basic methods, (a) automatically, running a user test using a software, (b) empirically, tests involving users, (c) formally, having exact procedure and formulas when testing, and (d) informally, based on some rules but relies on the skill of the evaluator (Nielsen, 1994c; Nielsen & Molich, 1990). Heuristic evaluation is considered an informal usability test (Nielsen, 1994c). This is because it is a flexible evaluation method that follows a set of heuristics while examining a website.

Heuristic evaluation is a usability inspection method in which evaluators examine a website according to predetermined principles. This method was selected for this project because it is known "…for optimizing workflows, improving user interface design, and understanding the overall level of usability of the website" (Kaushik, 2007, 58). Understanding the usability of the website is important to this study if the creation of the criteria are to reflect the best practices of digital museum information architecture. Additional methods that could have been used include cognitive walkthrough, heuristics walkthrough, and specific users studies. Cognitive walkthroughs use personas and tasks that would be performed on a website, for example, a digital museums "researcher" persona might have to find a specific image in the collection (Wharton et al., 1994). This method was not selected because it was not feasible for the timeline of the study and access to digital museums user profiles was not available. To create accurate and comprehensive personas and tasks, user studies on who and how people use digital museums needed to be completed (Wharton et al., 1994). It is also for those reasons that heuristic walkthroughs and user studies were not chosen for the study. Heuristic walkthroughs are an amalgamation of heuristic evaluation and cognitive walkthroughs, using a list of heuristics to evaluate a website according to personas that represent the target audience (Friess, 2015; Sears, 1997). Users studies in general can include focus groups, interviews, surveys, as well as usability inspection methods like heuristic evaluation and cognitive walkthrough. That would be an excellent place to start future research studies.

Understanding how digital museums use information architecture is a key part of this study. Jakob Nielsen and Rolf Molich (1990) the creators of the heuristic evaluation method, suggested that heuristic evaluation could determine the best design approach for a website (though with a different application of heuristics and evaluation techniques). This study will use a heuristic evaluation to see if there is a best design approach when creating the information architecture of digital museums. Some changes made to the traditional heuristic evaluation method include a new set of heuristics, multiple examinations of the select digital museum, and two report forms that evaluate IA principles and bugs/issues. Heuristic evaluations can find a large number of problems within the interface and are likely to find major issues, but this method also finds many low priority problems (Jacobson, Hertzum, & John, 1998; Nielson, 1992; Nielsen, 1995b). Any issues found will be reviewed again during data analysis, to make sure they are problems that should be dealt with.

The number of evaluators varies between studies, but the recommendation is three to five expert evaluators (those with experience in both usability tests and the domain) (Nielsen, 1994a; Nielsen & Molich, 1990; Novik & Hollingsed, 2007; Sauro, 2010). The researcher has worked as a usability coordinator for the past six years and has an art history undergraduate degree (so can be considered a double expert for this study). Evaluators examine the interface at least two times, first to familiarize themselves with the interface and second to conduct the heuristic evaluation. During this study, each digital museum interface will be examined three times, first to familiarize with the site (creating a contextual framework for each digital museum), then to examine the websites using the heuristics, and the third time to review decisions and the criteria suggested in the data analysis and findings stage of the study.

The study sample for the heuristic evaluation includes three examples for each classification of digital museum (brochure, learning, and content) (see Appendix E). The nonprobability sample method, purposive sampling, was used to select significant example(s) of the phenomenon being studied (Bickman & Rog, 2009; Leedy & Ormrod, 2010; Tashakkori & Teddlie, 2009). For clarification, probability sampling (not used for this study) is a method of sampling in which a random sample is selected from a sample frame that contains the entire population of whatever is being studied, this process ensures that every member of the population has an equal chance of being selected and is therefore a representative probability sample (Bickman & Rog, 2009; Leedy & Ormrod, 2010). A reliable and complete sample frame for digital museums did not exist, so a probability sample could not be chosen. Due to the type of sample selected, the results cannot be generalized to all digital museum information architecture design (Bickman & Rog, 2009). This is somewhat combated by the selection of different types of digital museums. According to Merriam and Tisdell (2016) "maximum [sic] variation in the sample, whether it be the sites for a study or the participants interviewed, allows for the possibility of greater range of application by readers or consumers of the research" (257). While a generalization of results cannot occur because of the sample selection, the reader may interpret the results for a similar study with similar context, content, and sample.

The digital museums were selected for a variety of reasons. All the digital museums are live websites, meaning that they have been updated recently (2016 or better) and maintained. That information is available either at the bottom of the website homepage or in the about page. These sites are most likely to have up-to-date information architecture, so the current uses of IA components can be evaluated. In order to narrow the focus of the study, mobile and/or tablet interfaces of digital museums were not evaluated (even though cross device/platform access is an important aspect of information architecture as discussed in Chapter 2). Metadata standards differ between languages; so digital museums from English speaking countries were selected (from Canada, United States, and United Kingdom). Each of the sites needed to have at least one example from all of the main IA categories; not having a search function was the most common exclusion. Finally, the size of the digital museum was a factor. Due to the time constraints for this thesis, a few of the possible samples were too large to adequately evaluate in the timeframe (e.g., the *Google Cultural Institute, Artstor,* or *The Metropolitan Museum of Art*).

The sampling criteria narrowed the number of digital museums selected for this study, but there were still thousands to choose from. The reasons for selecting the specific digital museums used in this study (see Appendix E) was because they met the sampling criteria, they had a large collection of images (meaning that there were sizable enough to evaluate but they were not so big that they could not be evaluated in the timeframe of the study), and they were interesting to the researcher (this is an example of author bias, see below for more information).

Heuristic evaluation data collection is built into the definition of the study. Each of the nine digital museums will be evaluated using the heuristics created in the content analysis study. The information collected during this process will be recorded down in a report format, outlining and providing an example of the heuristics found in the digital museums. The bugs/issues will be recorded in a Microsoft Excel form, this tool allows evaluators to justify if the issues are a problem (in an effort to prevent false positives) and provide additional context for the re-evaluation (Cockton & Woolrych, 2001). The bugs/issues form will have separate fields for a description, location, date, the heuristic it belongs to, and field notes made during the review (see Appendix D). The date is recorded because the examination takes place using live websites that can change very quickly. Once data collection is finished, data analysis will begin to create the information architecture criteria for select digital museums.

The data analysis for this study will compare and contrast the information architecture components. The information architecture principles discovered in each of the websites will be compared between the different types of digital museums (brochure, content, and learning museums). This will include descriptions of the principles as well as examples of how they were implemented in the digital museums. Once that initial comparison is complete, all of the digital museums evaluated will be compared together in order to create a list of information architecture criteria found in select digital museums. The criteria will be available for use by other researchers and contain examples, two items that Jakob Nielsen specified so that recommendations increase usability (Nielsen, 1994b). The creation of a list of information architecture criteria could provide consistency between websites, making the site easier to learn and more familiar for users.

During data analysis special attention will be paid to the bugs/issues that were discovered. These will be examined and retested to ensure that there are no false positives. False positives are issues recorded in a heuristic evaluation that are not actually problems (the subjective perspective of the user) (Cockton and Woolrych, 2001; Sauro, 2012a; Sauro, 2012b; Sauro, 2016). Heuristic evaluation results reveal bugs/issues found in the interface, but not how they can be fixed, however, because the issues have been associated with a heuristic it is easier to figure out a solution (Nielsen, 1994b). For any problem or bug found in the interface (that are not false positives), additional research will be conducted and solutions suggested.

Usability and Related Studies

Usability refers to how easy it is for users to use a website (no matter what skills or experience they have), specifically focusing on if the site is easy to learn (learnability), if users can learn to use it quickly (efficiency), if users remember how to use it when returning at a later date (memorability), if users encounter any errors when using it (errors), and if they are happy when using it (satisfaction) (Krug, 2014; Nielsen, 1994b; Nielsen, 2012). If those five characteristics align in a website you have a usable site. Determining the usability of a website involve extensive usability/user studies. Heuristic evaluation is just one example of a usability inspection method (generally volunteer participants evaluate the digital interface, not one researcher) (Nielsen & Mack 1994). Usability research is all about checking in with the actual intended/target audience of a site. You can design the greatest looking site on the web, but that does not matter if no one will use it. This is why research methods like interviews, focus groups, questionnaires, etc. are important tools when developing a websites design, particularly when creating the information architecture of a site (as discussed above). Usability is very important in digital museums. Lynne Teather (1998) wrote that usability tests needed to be conducted when designing digital museums, throughout the development. In the beginning the digital museums should establish whom their audience is, then conduct user testing continually throughout the design and production stages. Usability needs to be a priority for digital museums because they are so user focused. Here are a few usability examples related to evaluating digital museums.

Athanasis Karoulis, Stella Sylaiou, and Martin White (2006) discuss the evaluation of Augmented Representation of Cultural Objects (ARCO), a digital museum interface. There were two user study methods applied, questionnaires and a cognitive walkthrough. Museum curators (from the Victoria and Albert Museum in London, UK) answered the questionnaires, and museum visitors took part in the cognitive walkthrough (from the Aristotle University of Thessaloniki, Greece) (Karoulis et al., 2006). It was a quantitative expert–based study (meaning all the participants were an expert in their fields). The finding of the evaluation was that "… in complex interfaces, double experts (usability and domain experts) are inevitable for reliable and valid results" (Karoulis et al., 2006, 375). Meaning that double experts (as in this study) produce the best results in a usability study. However, when designing the interface/information architecture it should be the target audience who ultimately decides how to improve the usability of a site (but expert review is a useful starting point).

There is study that uses heuristic evaluation to examine the information architecture of a website. J. Parandjuk's (2010) used this method to evaluate the Publication of Archival Library and Museum Materials (PALMM) digital collection. Using information architecture guidelines put forth by Peter Morville and Louis Rosenfeld (2007), she evaluated the site to see if it adhered to those suggestions. Parandjuk determined that the PALMM digital collaborative is an excellent example of information architecture best practices, which can guide future website design (Parandjuk, 2010). This is a very relevant comparable study and uses the same methodology applied to a cultural website. The explanation and summary of the findings also use PALMM information architecture examples. This includes an examination of the vocabulary used by PALMM, which changed to reflect the intended audience of a specific collection. However, Parandjuk's (2010) heuristics were general questions (not disclosed to the reader), so they could not be used for additional research. The trustworthiness of the article is called into question as the author is employed by an organization working with the PALMM website, which is not articulated. The subject of author bias is an important consideration during all stages of research.

A German study used a heuristic evaluation, along with a think-aloud usability study to evaluate the Saarland museum website. Ilse Harms and Werner Schweibenz (2001) conducted the heuristic evaluation (following the *Heuristics for Web Communication*) and usability study using graduate students from the Saarland University as expert evaluators and volunteers as usability participants. General knowledge says that using both a usability inspection method and a usability study is the best method for evaluating a website, though this study states that "With respect to the cost-benefit ratio, in many cases the heuristic evaluation is sufficient to detect a reasonable number of minor and major usability problems"(Harms & Schweibenz, 2001, "Conclusion"). This article justifies why heuristic evaluation is a good choice for a study conducted by one evaluator (it finds a number of problems in the examined interfaces).

John Pallas and Anastasios Economides (2008) created their own quantitative digital museum evaluation method, "Museum's Sites Evaluation Framework (MUSEF)," and used it to examine 210 digital art museums from around the world (this evaluation was completed by John Pallas). The criteria used during the study had six categories (Content, Presentations, Usability, Interactivity & Feedback, E-services, and Technical), with subcategories organized underneath (each heading had a rating out of five) (Pallas & Economides, 2008). They discuss the results in statistics and came to the conclusion that North American museums had a better quality over the rest of the world, likely due to available resources (money, technical experts, etc.). This is important because it gave an example of alternate heuristics and methods used by one researcher to evaluate digital museums.

A proof-of-concept-study was completed in 2012. Three digital museums were examined using information architecture principles. One of each digital museum type (brochure, content and learning) was examined during a content analysis study (for both the information architecture literature and the digital museums – there was not a heuristic evaluation) (Sellmer, 2012). Results indicated that the select digital museums did follow some or most information architecture best practices, but not all. There could be improvements to the website's information architecture (fixing bugs). The pilot study showed that this is a valid and rich area to study, but the research design and methodology could be improved in order to provide further interpretation and validation of the data. This is often the case in pilot studies, they may take some time to do, but they let you know what does and does not work to help you solve your research questions (Leedy & Ormrod, 2010, 111). For this reason, new usability inspection methods were introduced.

The usability and related studies are similar (or the same) as the research proposed in this chapter. This project looks to overcome any shortcomings found in previous studies and/or utilizes aspects of their research. These past studies justify the need for this thesis because it will both improve the methodology design and introduce new heuristics that can be used for future research.

Validity and Reliability

To create a thorough study, special attention will be paid to the validity and reliability of the research. This will include a closer look at any biases present in the study. There are two classifications of validity – internal and external (or credibility and transferability, depending on the source), which are made up of best research practices. Summarized from multiple sources (Bickman & Rog, 2009; Leedy & Ormrod, 2010; Maxwell, 2009; Merriam & Tisdell, 2016; Roller & Lavrakas, 2015; Tracy, 2013), this list represents the different ways a study can improve the validity of the research and design:

Internal/Credibility validity:

- Readers need to be able to use the research and findings to draw their own conclusions.
- Show the readers the length of time spent with the data, so that any initial findings are confirmed or disregarded (if superfluous).
- Get feedback from other researchers.
- Be open to findings that provide evidence against your theory, and present examples to the readers.

- Compare the different steps of the study with other methods throughout the project – e.g., use two types of methodology for data collection to compare findings (triangulation).
- Use quantitative methods to support (or question) your qualitative findings (or vice versa).
- Present the research back to the study participants to see if they agree with the findings.

External/Transferability validity:

- Design and present the research to readers, so that they can apply the study to other contexts.
- Describe the data collected in as much detail as possible so that the reader can have a clear understanding of the study.
- Choose a representative sample for your study.
- Design your research so that the findings can be generalized (generalizability).
- Conduct the study in a "real-world" setting (e.g., not laboratories); findings will be more applicable to other real world studies.

This study will endeavor to follow the suggestions above if applicable.

The internal validity or credibility of the study can be found throughout this thesis. A great deal of time will be spent on the research, from data collection to data analysis. One example of this is the amount of time spent on the heuristic evaluations of the digital museums. Each digital museum is examined three times, the first time to become familiar with the website, the second time to evaluate the digital museum with the heuristics, and a third time to review the heuristic evaluation findings. This is one more time than suggested by the heuristic evaluation literature, but it ensured that there was a significant amount of time spent with each digital museum.

There will be only one methodology used for each of the research steps, so triangulation is not completely possible for this study. This is unfortunate, because it is an important aspect of internal validity. However, during the content analysis study, both qualitative and quantitative methods will be used. Using both methods allows for comparisons and supports interpretation with statistical facts. During data analysis if there were any findings that were unusual, these will be presented and discussed (with examples). A researcher is not always correct, so any contradictions to the theories or research questions will be examined.

External or transferability validity is achieved by presenting this study in rich detail. This includes, for example, the initial and edited coding frame for the content analysis study and the bug/issue evaluation form recorded during the examination(s). Researcher notes about data collection and analysis can be found on those and the content analysis sheets, so that the reader can understand the decisions made. The evaluation of the digital museums occur within the live version of the website (a real world setting). This does bring up a reliability problem because live sites can be unpredictable, but noting the dates that the bugs/issues were found will help deal with this problem.

For this study a representative and probability sample was not possible for either the digital museums or the information architecture literature. The actual number for either of those samples is unknown. Probability sampling would be the preferred method (selecting a random sample from a list of the complete study population, e.g., choosing every fifth book from a list of every art history textbook ever written), but not achievable for this study (Bickman & Rog, 2009; Leedy & Ormrod, 2010). The sample frame that came the closest to being reliable was a list of resources from the Information Architecture Institute, but it did not have any literature after 2015. The reason why not having a probability sample is important is because it means that this study is not generalizable. Generalizability means that you can apply the findings to every member that the sample represents. The best that this study will be able to do is offer readers an extrapolation, meaning that on a smaller scale "… the findings can apply to other situations under similar circumstances" (Merriam & Tisdell, 2016, 255).

Reliability for this study is harder to discuss, because reliability is about the consistence of the methodological instruments, and in qualitative studies the researcher is the instrument (Leedy, and Ormrod, 2010). The interpretation of data in a qualitative study means that findings can differ from person to person. To combat reliability issues, this study will have an "audit trail," meaning decisions are documented for the reader to judge, and decide if they would have reached the same conclusions (Merriam & Tisdell, 2016). The research planning, data collection, and data analysis documents are in appendices for the reader to review.

It is important that researchers consider the validity and the reliability of the content analysis study throughout the process. Due to the interpretive nature of a qualitative content analysis, these aspects of the study are more important than ever (Potter & Levine-Donnerstein, 1999). Consistency in the planning and a review of the coding frame is of the upmost importance (it is the foundation for data collection). There needs to be a pilot study because using the coding frame incorrectly threatens the study (Potter & Levine-Donnerstein, 1999), additionally, re-examining the coding frame multiple times can also reduce author bias (Schreier, 2014). One way to test the validity of the study is to refer back to the content analysis guidelines. For example, your categories need to be clear and precise, they also need to have at least one coded text within them, if they do not, then the pilot study and re-evaluation of the coding frame was done incorrectly and therefore is invalid (Schreier, 2014).

The reliability of the content analysis study is determined by the coder(s) interaction with the data. Called intracoder and intercoder, these terms classify how reliability can be determined in a study. Intracoder depends on the stability of the study across time. If the decisions made to the coding frame are retained after an initial retest of the rules (the pilot study), then the coding frame is stable for the main analysis (you can add more information) (Potter & Levine-Donnerstein, 1999). Intercoder is a far more complicated term. Intercoder has four reliability facets, (a) a calculation of the agreement between the study coders (coefficient calculated by methods like Krippendorff's alpha equation), (b) the trustworthiness of the sample, (c) an agreement between the coders about the coding frame and subsequent discoveries and (c) the retesting of any findings that appear unusual (Krippendorff, 2004; Potter & Levine-Donnerstein, 1999; Lacy, Watson, Riffe, & Lovejoy, 2015). This study only has one coder, so a coefficient cannot be established and a coding agreement cannot be reached (all the data produced in this study will be reviewed by two thesis advisors).

The article, *Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods*, by Gray and Salzman (1998) provides a deep insight into validity and reliability issues found in heuristic evaluations. The possible false positives are an intrinsic validity problem, which will be dealt with through a deep examination of the problem during the data collection and then another evaluation of the issue during data analysis. The authors of the article also describe issues with internal validity (setting of the evaluation, selection of and organization of participants into groups), and the external validity (claims that exceed the study scope) (Gray & Salzman, 1998). The third evaluation performed will help establish both the internal and external validity by spending a lot of time with the interfaces, which prevents superficial issues and allows for the presentation of rich data for the reader (Bickman & Rog, 2009; Leedy & Ormrod, 2010). While some of these issues are not applicable for the study (e.g., there is only one evaluator), the others are a priority. During the evaluation of the study, the same computer in the same room will be used for all of the three evaluations of the nine digital museums for a consistent setting. To make sure that the findings do not exceed the scope of the study, examples and rich description will be used to justify every solution and criteria suggested.

Author Bias

Another important factor in the validity of a project is the presence of author biases. Researcher biases are very important in a qualitative study, because their analysis and interpretation is an integral part of the process and should be valued (Connaway & Radford, 2017). However, if left unsaid, they can undermine the project. The educational background influenced the design of this study. When studying for an undergraduate degree in Art History, a lot of time was spent using digital museums (as a resource, for images, etc.). This experience with digital museums showed that information architecture of digital museums could be improved. For example, search systems could be poorly implemented or the organization of the website content could confuse the user. Yet there were wonderful experiences with digital museums as well. Preconceived notions do exist about the design of information architecture in digital museums, but personal history with digital museums is also the reason for this study.

Limitations

Every research project has limitations associated with it. The most important limitation for this study has to do with the design of the methodology. The structure of this study (first a content analysis then a heuristic evaluation) means that the findings from the content analysis determine the findings of the heuristic evaluation. As Anaganes et al. (2016) stated, "The choice of a heuristic set for a given evaluation changes the outcome of that evaluation" (586). This means that if there is a problem with the results of the content analysis, the research will be unreliable and invalid. To make sure this doesn't happen, the content analysis study (and indeed the entire project) will be done with all attention and consideration due to both the procedure and validity and reliability issues.

A limitation for the content analysis portion of the study is that there is only one coder for the data. Best practices say that there should at least be two or more (three or more would be better) so that analyses can be compared and a coefficient calculated (Lacy, Watson, Riffe, & Lovejoy, 2015). The only way that this study can reduce this limitation is with the presence of thesis advisors who will review the findings. Krippendorff (1989) discussed that the content analysis findings could not be generalized, because the data collected is from specific documents. Findings in the content analysis will have to be confirmed (using the coding frame) in another study.

A heuristic evaluation limitation that directly relates to this study is the use of only one evaluator (there should be at least three), but the researcher is a double expert in the field and that does produce better results. Multiple sources have described the limitations of having only a single evaluator, these include not finding very many problems, not being able to compare findings, and not being able to assign priority labels to the problems. (Jacobson, Hertzum, & John, 1998; Nielsen, 1992; Nielsen, 1994; Novik & Hollingsed, 2007; Sauro, 2010; Sauro, 2016; Woolrych et al., 2011). However, it is better to find some issues then none at all (Gould, & Lewis, 1985; Nielsen & Molich, 1990). Assigning severity ratings distinguish what is important to fix right away and what is of low priority, applied to bugs/issues by evaluators after coming to an agreement about the severity of issues (Andre, et al., 2003; Jacobson, Hertzum & John, 1998; Nielsen, 1995a; Nielsen, 1995b; Novik & Hollingsed, 2007; Woolrych et al., 2011). Unfortunately, for this study a single evaluator cannot apply severity labels because they would be unreliable (it would just be the subjective interpretation of a single person that could not be compared) (Nielsen, 1995a).

An additional limitation in heuristic evaluations is false positives. These are bugs/issues recorded in a heuristic evaluation that are not actually problems (just something that the evaluator through was a problem based on their subjective interpretation of a heuristic) (Cockton and Woolrych, 2001; Sauro, 2012a; Sauro, 2012b; Sauro, 2016). In order to combat this limitation there will be a third evaluation of the site to review the recorded bug/issues to make sure that they all represent real problems found in the evaluated digital museums.

Ethics

Ethical consideration was factored into every decision made throughout the study. No human participants were involved, so that is not an ethical concern for this study.

Chapter Summary

This qualitative study seeks to improve the information architecture of digital museums. It does so using two methods for data collection and data analysis, content analysis and heuristic evaluation. The content analysis study will complete an in-depth analysis of four general knowledge information architecture books, in order to determine a list of information architecture principles. This list will be used as the criteria for a heuristic evaluation (a usability inspection method). The heuristic evaluation will examine nine digital museums to determine what information architecture principles they use in their design and if there are any problems/bugs associated with it in the site. Findings will be presented in full detail to provide the user with as much information as possible, so that they can make an informed judgment of the findings. There are related studies that use the same or similar methods (including a pilot study), so this thesis will seek to improve upon and advance knowledge for select digital museums information architecture.

Chapter 4: Data Collection and Analysis

Introduction

Data collection and analysis for this thesis was completed in two parts. The first portion of the research was a content analysis study, which examined four general knowledge information architecture books. Data analysis of the content analysis results created the heuristics that would be used in the next section of the study. During the heuristic evaluation nine digital museums were examined (three of each type – brochure, content, and learning) using the heuristics created during the content analysis portion of the study. The heuristic evaluation examined the information architecture of the digital museums (both the structure of and any bugs associated with it). Once this portion of the study was finished, the results were compared and contrasted to view the information architecture criteria found in those digital museums.

Content Analysis

The content analysis pilot study was conducted in order to create data-driven subcategories, which were organized under predetermined main categories (organization, labeling, navigation, search, and vocabulary systems). This stage of the study was used to test and improve the coding frame.

The main content analysis study analyzed four general knowledge information architecture books (including the book used for the pilot study). These were coded according to the predetermined categories and subcategories (see below for further descriptions). Just like in the pilot study, each chapter was read and then coded according to the updated coding frame. This included disregarding irrelevant literature, making decisions about how to handle overlapping categories (while maintaining the mutual exclusive nature of the subcategories) and ensuring that every entry adhered to the formal unit of coding (no more than four sentences per code). Once data collection was completed, the data was reviewed and edited to make sure that everything was coded correctly and that all categories were mutually exclusive.

The initial data analysis for the main portion of this study examined the similarities and differences between the sources and paid close attention to how the coded data was broken down between the four sources. The relationships between the coded categories were also examined to determine how the coding entries intersected with one another. This included a closer look at the entries that had been coded more than once. Once the preliminary data analysis was completed, the list of heuristics (with definitions) was created, by reading through the coded text again and selecting the information that would be helpful for definitions (this included descriptive information, examples, components, and pros and cons). This information was them summarized/compiled into a list to be used as a reference document during the heuristic evaluation (see Appendix K). The list of definitions informed guidelines (list of questions to ask of the site during the evaluation) to be used in conjuncture with the definitions.

Pilot Study

The pilot study was conducted using the book, *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango. This book was chosen for the pilot study because it was widely considered to be the most comprehensive information architecture book. The *Institute of Information Architecture* described this book as "…the essential read for anyone looking to understand the principles and concepts involved in IA practices." (The Institute of Information Architecture, 2017, What is Information Architecture?). The text of each chapter was read and reviewed for information about information architecture principles. Any text determined relevant by the coding rules was excerpted, coded and organized into a Microsoft Excel spreadsheet. It was found that throughout the process the coding scheme was not restrictive enough in what should and should not be coded, resulting in numerous passages that did not specifically discuss information architecture principles. For example, in the latter chapters of the book they discussed "Sitemaps," but not the type found within a website, rather a research and design version that presented findings to the stakeholders of a website. For instance, this passage was coded:

> "Design is where you shape a high-level strategy into an information architecture, creating detailed sitemaps," (Rosenfeld et al., 2015, 314)

This was not an information architecture principle. Added to the coding frame, was a warning about the usage of similar or the same terms with different meanings. These needed to be closely examined (using the surrounding text for context).

Once the pilot study was completed, subcategories were created through subsumption. This was a data-driven process that created subcategories by checking if a category already existed, if it did not, create a new subcategory, if it did, subsume the coded entry it into the existing subcategory (Schreier, 2012). The list was created as new categories were encountered. For example, in the first chapter organization schemes were mentioned, so any information about or relating to organization schemes could then be subsumed. Using the information classified under organization schemes, subcategories were created (using the method of subsumption once again). An example of the final organization scheme hierarchy was: Organization System > Organization schemes > Exact organization scheme > Alphabetical schemes. Those categories were recorded as they were made in a Microsoft Word document (Appendix F). Many of the subcategories were left intentionally broad in order to collect as much information as possible about a subcategory and then further break that information down during analysis (in the main content analysis study). For example, under "Labeling systems," the subcategory “Labeling consistency” could refer to many different things. For example, was the grammar of all the labels correct? Did all the labels follow the same syntax? Had the same font? To make sure that this category covered all aspects it was left as a broad category. This allowed for further examination during data analysis (viewing relationships between coded text) in the main content analysis study.

Both the “Thesauri” and “classic thesaurus” categories were included. These seem to refer to the same category, but in this case “Thesauri” was used as a top-level category to organize other thesaurus information under. This category was also necessary to catch general information related to thesauri (that could not be classified into one of the lower more specific categories). For example, when thesauri were compared to the real world version within the sources, this discussed an information architecture principle (the thesauri), but not a particular version and in a more “general knowledge” metaphorical way (to help people understand what would otherwise be an invisible component of a website).

Once the full list of categories and subcategories was complete, miscellaneous categories were included as residual categories. Meaning that these contained items that did not fit elsewhere, which guaranteed that everything coded had a place to go (Schreier, 2014). Once data collection was completed for the main study, these categories were further analyzed.

The coding frame was updated using information collected during the pilot study (see Appendix G). Changes to the coding frame included descriptions for the new subcategories, changes to the coding rules, and additional classification of relevant and irrelevant information. The coding rules were updated to include further information about what to do when encountering overlapping information. Keeping in mind that the first priority in these instances was maintaining the mutually exclusive nature of the categories. “Hybrid” subcategories were discovered, which is text that discussed two or more subcategories used together to organize information. Coders needed to keep in mind that overlapping subcategories may belong in a hybrid category, depending on what was being discussed and the context around it. The coding frame now contained examples, to help clarify the coding rules. An example of an overlapping excerpt, taken from the pilot study was:

> “Or, you might ignore synonym rings for initial searches but provide the option to ‘expand your search to include related terms’ if there were few or no results.” (Rosenfeld et al., 2015, 275)

You could code this under Synonym rings and Repeating/new search. This would be coded twice.

Additional coding rules were added about subcategories that were similar or (in some cases) the same, but were categorized under two different main categories. For example, contextual links were found under both labeling and navigation systems. For the main content analysis study, the coder needed to remember that those categories could exist in multiple places (as long as they remained mutually exclusive) and used context to determine where they belong.

The list of irrelevant material increased because too much information was being coded that did not directly relate to the research question (like building architecture information). For example, prior to the pilot study the coding frame identified that information in tables and images would not be coded, after the pilot study it was added that the description of both tables and images (commonly found in italics directly beneath them) would also not be coded. There were three reasons for this decision (a) the actual surrounding text almost always covered/discussed the information represented by the images and tables, (b) the text of the literature was the focus for the study, and (c) including those features would not be feasible in the time frame of the study. Additionally, the decision was made to not code quotes from secondary sources. Only primary information would be coded maintain ethical standards (not copying a quote from someone else's citation).

Changes were also made to the coding form (see Appendix H). Once the pilot study was completed, it was discovered that the form used did not provide opportunities for statistical analysis because codes were textual, not numerical. This was changed so that the codes were now numbers from 1-115, applied by how they were organized in the categories and subcategories list (see Appendix F). For example, Organization System would be 1 because it appeared first on the list and the miscellaneous subcategory in Vocabulary systems would be 115 because it was last. All categories and subcategories were assigned a code because each level was representative of information architecture principles. For example, when discussing labeling systems some of the principle only talked about "Textual labels," and could not be classified into "Headings," "Contextual links," etc. So that text excerpt would be coded with the number 24 (the "Textual label" code). This happened throughout the pilot study because information about information architecture principles spanned from broad descriptions to specific examples.

There would also be numerical codes for the source material (numbered 1-4) and unit IDs given to every coded unit (1-*n*). Unit IDs would be given to each coding unit, so that the Excel spreadsheet could always be organized back to how the list was originally coded. For example, when looking at what search features have been coded (in all sources) one could organize the Excel spreadsheet so that those categories were arranged together. Once you had manipulated the sheet, if you want to return it to how it was originally coded one would need to sort the unit IDs by "ascending" order, returning the list to normal. Now you could begin to analyze the data again. If a unit of text was coded more than once (this was acceptable as long as they were not under the same main category) they had different unit IDs because they would be treated as separate units of coding.

The main content analysis study focused on establishing the relationships and importance of the categories and subcategories through a close examination of four general-knowledge information architecture books. The categories and subcategories created during the pilot study formed the foundation of the heuristics used in the heuristic evaluation, but the text coded within them will be closely examined in order to create the final list of heuristics. This included any information in the miscellaneous categories. The list of heuristics will also have descriptive information (drawn for the content analysis study). This included definitions for each term, examples of the heuristic in real world practice, pros and cons elements of the heuristics, and identifying components (e.g., how can you recognize a database-oriented organization structure in a website).

Data Collection

Data collection for the main content analysis study worked through the four select general knowledge information architecture books. Text that contained information about or relating to information architecture principles was excerpted and coded. If there was a subjective decision made about the coding, it was noted it in the memo section of the coding form. Once all four sources were analyzed, the coding form was edited. The collected data was reviewed to ensure that every unit was mutually exclusive, followed the rules of the coding frame, and that they were all coded accurately.

Data collection began with *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella (coded 1). This book differed from the sample used in the pilot study because it was a less technical in the descriptions of the information architecture principles, so there was more interpretation needed to code the excerpted text. The book still discussed most of the categories and subcategories (and added a few more that were sorted into the residual categories). Christina Wodtke and Austin Govella (2009) were the publishers of the online blog *Boxes and Arrows* and that site was used as an example throughout the book. The final chapter discussed the information architecture redesign of *Boxes and Arrows*,[2] outlining the decisions and procedures made by the publisher, editors, and information architect. This was an excellent example of how information architecture principles were implemented in real world sites (which you could then go online to view).

The second source coded was *A Practical Guide to Information Architecture* by Donna Spencer (coded 2). A lot of attention was given to research, users, and working with clients, but most of the information architecture principles were covered in this book. A real world example was also used throughout the book, the website for the user experience (UX) conference held in Australia (Donna Spencer was the co-organizer).[3] For example, this website had an upside down "L" shaped navigation system, which was discussed in the navigation chapter (Spencer, 2011). This book provided a great look at how to build information architecture into websites and how information architects work with clients to explain and implement IA principles.

The third book coded was *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango (coded 3). This was the sample used for the pilot study. The bulk of the book discussed the main categories (organization, labeling, navigation, search, and vocabulary systems) and it had the most entries in the final coding form. This book examined information architecture principles in detail, and because it was used for the pilot study it resulted in very few text excerpts coded into residual categories (across all sources). The authors also discussed both the contextual framework around information architecture (the importance of audience, content, and context), as well as research and design. This was an excellent source for the content analysis study

The fourth and final book coded was *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert (coded 4). This was the shortest of the four samples (at 180 pages) and it emphasized "general knowledge" over specifics. Abby Covert's (2014) goal in this book was to provide a broad outline of information architecture, so only a few categories were written about in depth. This included information about taxonomies, which could refer to

---

[2] http://boxesandarrows.com
[3] http://www.uxaustralia.com.au

classification schemes, organization systems or hierarchies. Meaning that this book required quite a bit of interpretation when assigning codes (e.g., does this belong in classification scheme or under organization scheme). Like the other three books, discussions about users, content, context, research and design were present throughout the chapters.

Throughout data collection, the coding frame was followed and referred back to any time there was a question about how a unit should be coded (see Appendix G). For example, in the first source (*Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella) the terms "Pogosticking" and "Crabwalking" were introduced. These were determined not to be principles but rather descriptions of user behaviour, so the text would only be coded if there was an information architecture principle discussed. All chapters were coded, but the "supporting" sections were not (preference, table of contents, references, etc.) because they did not contain the text of the literature. Principles would not be found in those sections other than the occasional out of context mentions (e.g. the index). While coding the sample chapters, each excerpt of text ranged from one word to four sentences. The formal unit of coding was decided on so (a) sentences that contain more than one principle can be separated and (b) so that each principle can be specifically coded without too much unnecessary information (read: non principle related) surrounding it. The only reason that the full four sentences were coded was if they all referred to one principle. For each text excerpt coded, the page number for that section was added. The readers could look up the coded text in the books (using the source code and the page number) to read the text in context if they were so interested. This also represented a citation for the coded text.

The relevant and irrelevant materials rules were followed so that only information directly related to the research question was coded. This set the parameters of the study. The main focus of this study was the information architecture principles, so during data collection this was carefully considered. For example, Donna Spencer discussed and explained information architecture throughout the beginning of her book. These were generally not coded because they did not mention the individual principles. Here is an example:

> "But good information architecture can do more than just help people find object and information. It can empower people by making it easier for them to learn and make better decisions." (Spencer, 2011, 4)

This quote described information architecture and how it helped users, but this was not coded because this passage of text did not discuss an information architecture principle.

Any materials that did not represent information architecture principles were considered outside the scope of the study. This included information about graphic design, building architecture, website users, research design, wireframes, etc. The coding frame outlined these decisions. For example, in this quote *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert, she discussed how users impact information architecture. For example:

> "While we can arrange things with the intent to communicate certain information, we can't actually make information. Our users do that for us." (Covert, 2014, 13)

Users were at the center of information architecture, but they were not information architecture principles nor was one discussed in this excerpt, so it was not coded.

Tables, images, and figure descriptions were also determined to be outside the scope of the study. As mentioned above, these were excluded because of the timeframe of the study, the literature generally discussed them in detail elsewhere, and because the text was the focus of the study. For example, this next quote was a description for "*Figure 9-13. A Yellow Pages search doesn't force us to click through for a phone number*" (Rosenfeld et al., 2015, 235). It was not coded even though it discussed the "information displayed for retrieved items" (code 59). However, the text around this image and description was coded:

> "Users of phone directories, for example, want phone numbers first and foremost. So it makes sense to show them the information from the phone number field in the result itself, as opposed to forcing them to click through to another document to find this information." (Rosenfeld et al., 2015, 235-236).

This example showed that because tables, images, and figure descriptions were not coded, some examples of principles were missed. But these were generally described with more detail (and context) in the actual text of the book.

The coding rules were followed as data collection progressed (as described in the coding frame, see Appendix G). This included how to code each unit of text (from 1-115), what to do if there was overlapping categories in a passage of text, and how to deal with the ambiguity of terms. When handling overlapping categories in a unit of text there were three different options, (a) break the sentence apart (especially useful for maintaining the mutually exclusive nature of the categories), (b) code it under both subcategories (only allowed if the subcategories were organized under different main categories) and (c) look to see if they belonged in a "Hybrid" subcategory (which was when two subcategories were combined into one principle). For example, in *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella, this sentence needed to be broken up: "Local navigation often appears "below" the global navigation." (Wodtke & Govella, 2009, 197). This sentence could not be coded together because then the subcategories would not remain mutually exclusive (this sentence needed to be coded in both "Local" and "Global" navigation subcategories). The sentence was separated into two different units of text (after "below"). Dealing with overlapping categories was a challenge throughout the coding process because information architecture categories are closely connected.

There were different terms used for the categories that needed to be considered. For example, Abby Covert called organization systems by the term taxonomies. Close attention needed to be paid to the context surrounding these principles to determine where they should be coded. For example, this sentence: "Structural methods for organization and classification are called taxonomy" (Covert, 2014, 101). The surrounding text was read and it was determined that "structural methods for organization" referred to organization structures (that one was pretty easy to logic out). That "and classifications" belonged in organization schemes. And finally, "are called taxonomies," meant organization systems (under which organization structures and organization schemes were classified). This sentence ended up being coded into three different units of text, representing organization structures, organization schemes, and organization system. This was not the only example of ambiguous terms encountered when coding. For example, contextual links could refer to either a type of navigation or a type of labeling. The excerpted text was read and surrounding context considered in order to determine where it should be classified. For example,

> "Related links are a very common type of navigation item. As the name suggests, they allow access to content that's related to the current page. They allow you to highlight content you'd like to expose to people, and help users to find things they may be looking for." (Spencer, 2011, 267)

This could be coded into either contextual link labels ("Related links" was a common label used for contextual links) or into contextual navigation. It was determined that this should be coded into contextual navigation because, navigation was the main topic and the description discussed how you used those links (not how you read/interpreted them).

Once data collection was completed, the final coding form was reviewed to make sure that everything was coded correctly. For example, there were three mentions of the "shopping cart," commonly found in the top right corner of shopping websites. These were originally coded into a residual navigation category, but on closer inspection it was determined that these should actually be coded into iconic labels. For example, "…is accompanied by a fairly standard bag icon that implies 'shopping cart'" (Rosenfeld, et al., 2015, 136). This should have been categorized into icon labels to begin with, distance and a second look helped catch those problems. This review also determined that all categories contained at least one coded entry (the lowest having only one coded entry). For example, phonetic tools (under search query builders) had only one coded entry – a description from source 3. All 2516 units of coding were reviewed and once editing was finished, data analysis began.

Data Analysis

Analyzing the data for this study mainly focused on creating the list (and definitions) of information architecture principles for the heuristic evaluation. Prior to that, a preliminary analysis of that data was conducted. This looked at the relationships between sources and between the coded text excerpts. Questions this section answered included: What were the similarities and differences between the sources? What were the relationships between the coding? What text excerpts were coded more than once? These questions included qualifying quantitative information (when applicable).

The four sources used in the study (discussed in detail above) ranged from general knowledge to detailed particulars when discussing information architecture. For example, *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert discussed information more abstractly, using real world examples (like how grocery stores use organization systems – content was organized by aisles, grouped by similar items, and arranged on the shelf (Covert, 2014, 12)) instead of explicit discussions about the information architecture principles. This resulted in only 72 excerpts of text coded or 2.86% of all coded text (for statistical tables see Appendix I). Though it should be noted that this was also the shortest of the four sources (it was a small book and only had 180 pages). The most detailed source was *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango. It had the most coded text (by far) at 1481 or 58.86% of the total coded entries. This source covered all the categories and subcategories (which was to be expected because it was used for the pilot study), and in some cases it was the only source found within a subcategory. For example, the subcategory in the search section, "Indexing for recent content," only had text excerpted from *Information Architecture for the Web and Beyond*.

The other two sources, *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella and *A Practical Guide to Information Architecture* by Donna Spencer were almost equal in the specificity they used to discuss information architecture (Wodtke & Govella had 579 coded entries or 23.01% of the total entries and Spencer had 384 coded entries or 15.26%). Donna Spencer's book had less coded text because of the amount of time the book spent discussing client relations (this book is great for an information architect at the beginning of their career) and the use of secondary sources. *A Practical Guide to Information Architecture* had real world information architects share their experiences working with clients or how they implemented information architecture. While this was a great resource for readers, that information was not coded because of the coding rules (no secondary sources were allowed to be coded due to ethical reasons). These two sources used different terminology than the pilot study source (number 3). For example, they labeled structural metadata as intrinsic metadata (metadata about the items composition). The different terminology was noted when analyzing the data. For example, "Source 1 and 2 labeled structural metadata as "Intrinsic" metadata. For the organization into categories, they were one and the same" (under code 107).

Information Architecture: Blueprints for the Web by Christina Wodtke and Austin Govella (source 1) was responsible for the most information categorized into miscellaneous categories, especially within navigation systems (which made up 45.68% of all miscellaneous categories coded). This included the navigation items: utility navigation, control panels, toolbars, and pagination navigation. For example, this was coded (from source 1) “Amazon has a nice design for pagination. It tells you what page you’re on, offers links to the previous and next page, as well as links to specific pages” (Wodtke & Govella, 2009, 215, code 505). This did not belong under a pre-existing subcategory so it was coded under Navigation system > Miscellaneous. Upon analysis of the data it was noted that this type of navigation actually had been discussed in a passage of text from the pilot study source (3), vaguely:

> “Also consider providing a results navigation system to help them move through the results [search results]…. Reuters provides such a navigation system, displaying the total number of results and enabling users to move through the results set 10 at a time” (Rosenfeld et al., 2015, 238, code 1928)

This quote was overlooked in the pilot study, but this type of scenario was accounted for and the reason the residual categories were included in the coding frame (see Appendix G).

Despite the differences between the sources, each one had information coded beneath all the top-level categories (organization, labeling, navigation, search, and vocabulary systems – see Table 1). However, in the case of the second source (*A Practical Guide to Information Architecture* by Donna Spencer) and the fourth (*How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert), they only had three text excerpts (in total) coded under search system. Donna Spencer only mentioned the search system twice (it represented .52% of the coded entries for that source), and these were both so general they ended up being classified under the highest category, “Search system.” For example, “The best design solutions for this behaviour are search and A-Z indexes” (Spencer, 2011, 97, code 590). This was coded under both A-Z indexes and search systems. The one mention of search in *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert discussed sorting information (this was applied to sorting search results). The text excerpt was “Sorting is the act of arranging content according to established rules” (Covert, 2014, 103, coded 2469). For a full statistical breakdown of how each category (organization, labeling, navigation, search, and vocabulary) was coded by the sources, see Appendix I.

The category with the most coded content was navigation systems (25%), which was also where the miscellaneous categories had the largest number of coded text (see Table 1). Organization and vocabulary systems closely followed navigation systems in frequency of coding (23% each). These two categories were consistently coded across all the sources. Source 4 even discussed both of these systems multiple times, (though sometimes the terminology differed) despite the fact that it only had 72 entries in total. For example, "A controlled vocabulary is an organized list of terms, phrases, and concepts intended to help someone navigate a specific context" (Covert, 2014, 74, code 2453). This was only coded under controlled vocabulary even though it has the term "navigate" in the coded text. For a full statistical breakdown of the coded categories see Table 1.

Name	Percentage
Organization system	23.25%
Labeling system	11.09%
Navigation system	25.28%
Search system	17.89%
Vocabulary system	22.50%
Total	100.00%

Table 1. Distribution of Information Architecture Principles. This table shows the percentage of coded information that belongs in each category (for all sources).

Search systems contained 18% of the coded items. Understandable when you consider that search was only really discussed in two sources (1 and 3), and the bulk of that coded information came from *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango (source 3). This was where information about query builders was found, and source 3 was the only one that mentioned search zones. On the other hand there was almost no information about vertical search in source 3, but it was covered in source 1 (and so it was coded into the search miscellaneous category).

Labeling systems rounded out the coding distributions with 11% of the coded entries respectively. It was interesting that labeling was the category with the lowest amount of coded information despite the fact that it was an important component of information architecture and was found across all categories (e.g. index terms include descriptive metadata and global and local navigation were formats for displaying navigation labels). This revealed a problem in the study because on further analysis of the data, labels were not coded every time they were mentioned. Unfortunately, the other categories discussed in an excerpted text overshadowed any mention of labels. Meaning that they would only be coded when they were the more overt principle discussed. For example,

> "Allrecipes mixes several kinds of classifications. This means some bread recipes can be in more than one category. For example, you might find a hot cross bun recipe under Breakfast Pastries, Holiday Breads, Yeast Breads, Fruit Breads, and Rolls and Buns. This is okay. It may make purists itch, but it gets people to the bread recipe they need." (Wodtke & Govella, 2009, 55, code 73)

This excerpt was coded only under faceted classification, even though it referenced index term labels. This happened throughout the study and was a potential issue. Though one could argue that every time information regarding labels was discussed abstractly or indirectly they were not the topic of the excerpted text (and this fell into overlapping rules in the coding frame, specifically "The main topic is generally the more specific principle discussed" – see Appendix G).

The difficulty of coding information architecture principles was that the categories were all closely connected to one another (as seen above). For example, you could determine the organization scheme(s) of a website by looking at the labels assigned to the global navigation, and those labels were often determined by the sites controlled vocabulary, which also dictated what search results were returned. This example of overlapping information would had been okay because all those principles could be categorized under different categories (organization, labeling, navigation, search, and vocabulary systems) and the mutually exclusive nature of the categories would had been maintained. However, this could be difficult if the items belonged under one category. For example, a thesaurus was part of a website's controlled vocabulary and those two terms were often mentioned together in the same sentence (along with other vocabulary subcategories). For example, "With this look at thesauri, controlled vocabularies, and metadata, we conclude the 'basic principles' part of the book" (Rosenfeld et al., 2015, 309 codes 2364-2366). This example was split up into three coded entries. There were many more examples of different sentences split apart into one or two word sections to make sure that the categories remain mutually exclusive (not just under the vocabulary system). Global and local navigation were discussed together almost constantly (which was natural considering one depended on the other), so there were often coded text with just the word "global" or "local."

By looking at the categories coded more than once, the connections between categories and subcategories were visible. There are 75 entries coded more than once, and three of those were coded three times. For example, the quote "Another thing I've experienced is the difference between using a map to display information and using a map to navigate to it" (Wodtke & Govella, 2009, 162, code 383-384) was coded twice, once under geographic organization scheme and once under advanced navigation > visualization. Or this quote "This is the only verb in the global navigation, a potential source of confusion for users who may read it as leading to information about a physical 'shop'" (Rosenfeld et al., 2015, 139, codes 1336-1338), which was coded three times. This was coded under global navigation; task oriented organization scheme, and labeling consistency. This coded text excerpt not only showed a few of the relationships between information architecture principles, but it also referred back to the issue of not always coding label information when it appeared abstractly with other codes. This inconsistency needs to be addressed in future studies. Information that was often coded more than once included metadata, types of labels, A-Z indexes (especially in source 2), and hierarchy structure.

Once the first analysis of the data was completed for this study, a third review was conducted in order to create the heuristics for the heuristic evaluation. This involved creating a list of information architecture principles defined with the excerpted text and coded data from the four sources. To create the criteria for the next stage of the study, data was organized by code (ascending to descending), so that all categories and subcategories were organized together. This was printed out, because it was preferred method of analyzing the data. Then the coded data was reread and notes were added, highlighting the chunks of text that could be used to define the heuristic, examples of the heuristic, different components, and if there were pros and cons for the principle. These were written on the printed list of coded data (see Appendix J). Once the category or subcategory was reviewed, the coded text that would help identify those principles in real world websites was noted down, the description of the heuristic began. It is important to note that not all coded text excerpts were used to create these descriptions.

Findings

With the identification of coded text that could be used for the definitions – examples, components, and pros and cons for each category and subcategory, the final list of heuristics was created. This list of defined heuristics (outlined in more detail below) and the guidelines (a list of questions used during the evaluation to help clarify the information architecture principles the site utilized) were the heuristics used in the heuristic evaluation.

An example of defining a heuristic could be seen when creating the description for the organization structure "Social Tagging." There were 21 entries coded under this subcategory (code 19), which ranged from real world examples to two-word entries. Of the 21 entries coded, five were classified as useful for the definition of this subcategory (only three ended up being used). For example, these two quotes "Each item in the site is 'tagged' with keywords, and those keywords are used to provide access to the content" (Spencer, 2011, 205, code 789) and "The tagging may be done by the original authors of the content, the readers or by some central authority (such as the web team)" (Spencer, 2011, 205, code 790) were combined and summarized to form the first sentence of the definition. Those two sentences became, "This organization structure leverage tags (user or expert created) to provide access to content (789, 790)." When creating these heuristics, the Unit ID numbers were added as a reference and citation for the information used in the heuristic descriptions. There were two coded entries determined to be examples of this organization structure. They were added those to the definitions (organized beneath the heading "Examples:"). There were also two "pro" elements for this subcategory that were also added to the final description.

This process was repeated for every heuristics definition. An additional example would be defining the query builder, spellcheckers (under search systems). There were nine different coded entries under spellchecker. Of the nine coded entries, three were determined to be good examples for the definition (all from sources 3, *Information Architecture for the Web and Beyond*) and four entries were determined to be examples of this heuristic (all from source 1, *Information Architecture: Blueprints for the Web*). This coded entry "Yahoo!, however, recognizes the wide variety of spelling humans manage to invent, although ‘chedder’ works rather well, they also prompt you to try ‘cheddar’” (Wodtke & Govella, 2009, 81, code 173) was classified as an example for the final definition. For the full definition see Appendix K.

This list of heuristics included the organization and summary of text coded under the miscellaneous categories. These were examined in order to make sure the list of heuristics was as complete and extensive as possible. Of all the items coded, only 103 entries were classified into the miscellaneous categories (4.09% of the total data coded). Despite not having that many items coded in to the miscellaneous categories, there were items that needed to be added to the final list of heuristics. Those subcategories were: Format and Organizational (business) schemes (organization schemes), Linear pattern structure (organization structure), Utility navigation, Control panels (supplemental navigation), Toolbars (supplemental navigation), Pagination (supplemental navigation), Relevance ranking (ranking search results), and Vertical searching (search systems). When these subcategories were defined, the miscellaneous category they came from was noted in the heading of those definitions.

The final list of heuristics (see Appendix K) outlined all the information architecture principles for the heuristic evaluation. This included a closer look at “types” or subsets of the principles that revealed themselves during the content analysis study. For example, there was a hybrid organization structure subcategory, but *A Practical Guide to Information Architecture* by Donna Spencer (source 2) introduced different “types” of hybrid organization structures. In the final list of heuristics these were added with definitions and examples when available. For example, there was the common hierarchy/database structure, the hub and spoke pattern structure (the main page was the “hub” of the site and users moved into different “spokes” and then back to the “hub”), or the subsite structure (multiple sites held together by the homepage and website design). The final list of heuristics contained 119 principles.

The list of heuristics informed the evaluation guide, which was a list of questions asked of the nine digital museum websites during the heuristic evaluation (see Appendix L). This evaluation guide went through each category and turned the heuristics into questions. When examining the organization systems the guidelines asked not only what organization schemes and structures the website may have, but also where they are located and how the websites utilize them. In the labeling system, it was asked what type of labels can be found in the website (contextual, headers, icon, etc.) and if there are any consistency problems. The navigation system guidelines were what navigation features were used and where, when, why, and how they were presented in the website. The search system guidelines were more complex (to represent the multiple heuristics found under this category in the content analysis study). For example, when evaluating a digital museum's search system, the digital museum was reviewed in order to answer what the search interface looked like, was there an advanced search interface, how were the search results presented (format and information provided), were the search results ranked or sorted (or both), etc. The guidelines for the vocabulary system were specific because this is the most difficult section of a website to identify. In the guidelines, metadata, controlled vocabulary, thesaurus, semantic relationships, and faceted classification categories were outlined. For example, the thesaurus guidelines asked if the preferred terms could be identified, how the thesaurus was structured, and if it followed standards. This also included suggestions to look at indexes/keyword metadata to determine the thesis structure. Throughout the guidelines it was noted that examples and clarifying information should be added at every stage of the evaluation for contextual information and to make data analysis easier. This guide was used throughout the heuristic evaluation in conjunction with the list of heuristic definitions to provide the user with as much detail as possible to identify the information architecture principles.

Heuristic Evaluation

The heuristic evaluation examined the information architecture of nine digital museums:

- The Art Gallery of Ontario
- The National Portrait Gallery
- The Cleveland Museum of Art
- Art UK
- WikiArt
- Discover Islamic Art
- SHOW.ME

- Smithsonian Learning Lab
- Web Gallery of Art

This involved noting the principles they contained as well as any bug/issues encountered while inspecting the sites (related to information architecture principles). Reports were written during data collection following the evaluation guideline created from the findings of the content analysis study. These reports provided descriptions and examples of the information architecture principles each digital museum contained (see Appendix M for an example). The list of bugs/issues were recorded in a Microsoft Excel spreadsheet and included information like where the bug was found and in what museum, a description of the issue, and what information architecture principle it was associated with (including subcategory codes from the content analysis study).

Data Collection

The heuristic evaluation examined the select digital museums following a set of heuristics. These heuristics were created during the previously discussed content analysis study and were comprised of the evaluation guidelines and heuristic descriptions (see Appendices J and K). Roughly twenty-eight hours were spent evaluating each museum (some took longer if the site was more complex or had a confusing information architecture design). The first step during the heuristic evaluation was to become familiar with the site; navigating through it and creating a contextual framework for the digital museum being explored accomplished this. The contextual framework included information about the context of the site (the organization (business) plan, the purpose of the site, etc.), the audience of the site (who used it, how was this apparent in the site design, if there was an audience-oriented scheme, etc.), and the content (what did this museum present for the audience, this did not just refer to images, but also learning resources for example) (Covert, 2014; Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). The idea for the contextual framework was encountered numerous times throughout the literature review and content analysis study (though the information was not coded because it was not related to information architecture principles), so it was decided that creating a contextual framework for each digital museum would be an excellent way to become familiar with the websites.

Once the initial review and contextual framework was completed, the heuristic evaluation began. The first step of the heuristic evaluation involved going through each page of the website, noting down the page/heading title, the organization structure, how information was organized (was there an index, if the content was organized chronologically, geographically, etc.) and if there were any navigation features. Once there was an understanding of the websites organization (structure and scheme), a diagram was made using the information collected (see Figure 1). This diagram was double checked for accuracy because it was often referred back to for information about the organization and labeling of the site.

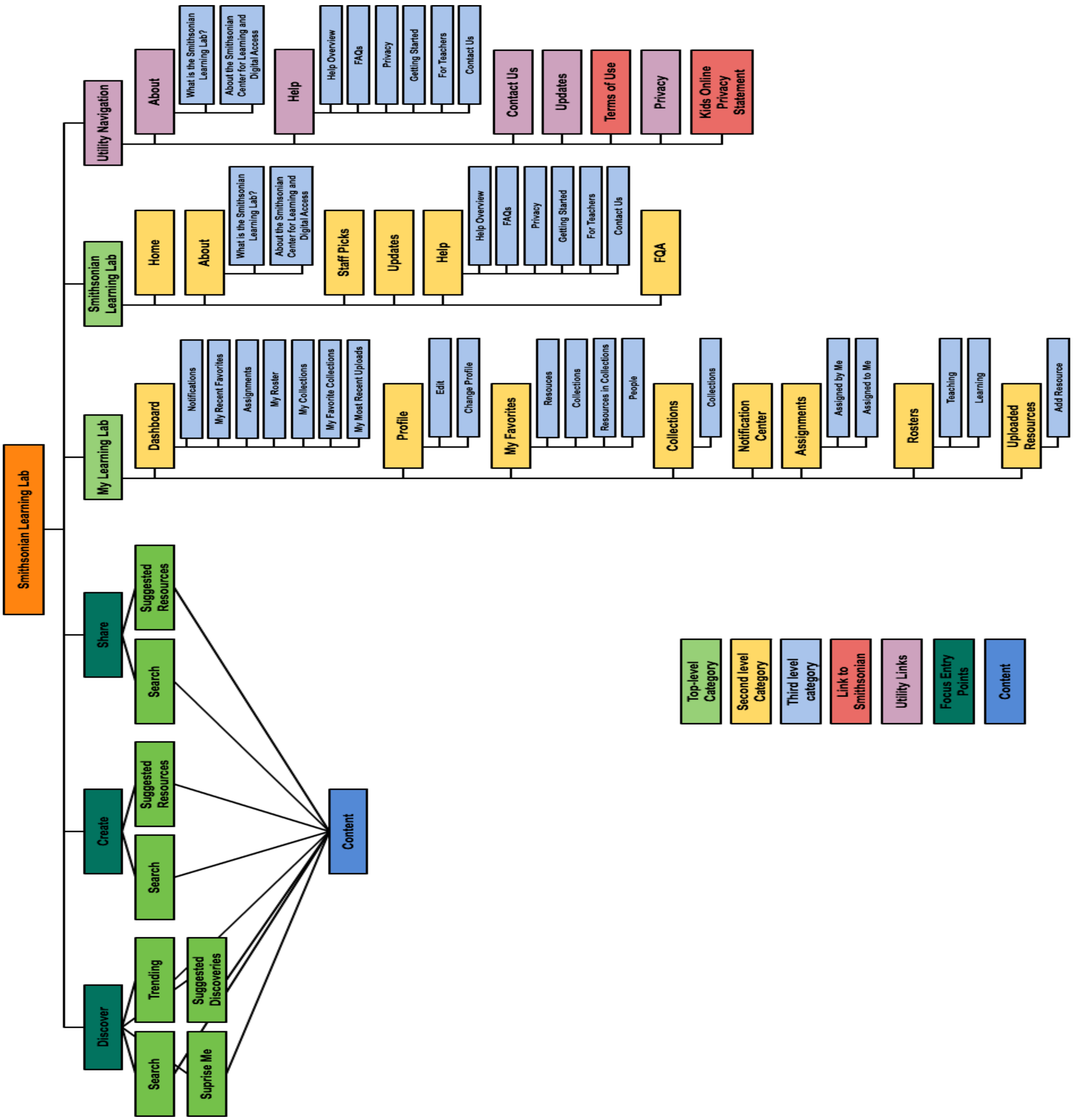


Figure 1. Diagram of the Smithsonian Learning Lab Digital Museum. This type of diagram made the top three levels of the organization system and the labeling system of the site visible.

Once the diagram was completed, observations made during the heuristic evaluation were written down in a report format. This involved going through the notes and diagram, and writing down the organization schemes (by level, e.g. top-level, second level, etc.), the organization structure, and the types of labels/labeling consistency issues found in the site. For example, when evaluating the ART UK website, the diagram revealed that the top two levels of the site follow topical (About, Artworks, Artists, Stories, etc.), task-oriented (Discover, Participate, Visit, Become an Artwork Detective, Shop, etc.), and Format (Blog, Art prints, Books, etc.) organization schemes. The term "level" refers to the way a user moves through the site, the top-level would be global navigation labels, the second level would be local navigation, and the third would be sub local navigation labels (etc.).

Once that section was completed, the navigation features available in the site were identified (what, where, when, etc.). This section was structured according to the list of heuristic descriptions from the content analysis study (global, local, contextual/hypertext, breadcrumb, utility, supplemental, and advanced navigation features). These descriptions included examples and images to enrich the information. For example, in Discover Islamic Art (a content museum), the museum report explained the location of the local navigation - within the "Exhibition," "Artistic Introduction," "Learn with MWNF," and "My Collection" global navigation categories. And then provided examples (e.g., within the "Learn with MWNF" subsite there was a visualization navigation feature).

After the evaluation of the digital museum's navigation was completed, the site's search system was evaluated. The guideline questions were answered for each type of search system available on the site (this ranged from nine different systems to only one). Test searches were conducted to evaluate the search system, writing down information that related to the information architecture search principles. For example, when evaluating the main search system in the National Portrait Gallery there was an advanced search feature. This was described in detail, noting the different sections that users could enter text into, "Person" and "Portrait." And you could select a role (artist or sitter), enter a profession, select a professional category, living/deceased, etc. Images were added to provide additional detail about what was being evaluated.

The final stages of evaluation looked at the vocabulary system for the site. These were the most difficult information architecture principles to identify because much of the vocabulary system was in the "background" of the site and needed to be deduced by the evaluator. Indexes, glossaries, search systems, faceted browsing, and metadata (which was accessible) helped determine the controlled vocabulary. For example, the faceted browsing or search filters could help determine if there were any classification schemes present in the site (having preferred predetermined terms, which the user used to browse information with). The index, glossary, or metadata was used (if possible) to ascertain the thesaurus of the digital museum. These terms were compared to thesaurus standards (Art and Architecture by Getty, the Library of Congress Subject Headings, etc.) and if they did not follow those standards, an attempt was made to structure a thesaurus from the terms available (to provide readers with a sense of the sites terminology). This section finished by describing the semantic relationships between terms and the faceted classification system (if present), according to the heuristics and guidelines.

All stages of this evaluation were done following the heuristics and guidelines created in the content analysis study.

Data collection for the heuristic evaluation began with the three digital museums classified as "brochure" museums. This included the Art Gallery of Ontario, the National Portrait Gallery, and the Cleveland Museum of Art (they were examined in that order). These were the digital representations of physical museums. Then the Art UK, WikiArt, and Discover Islamic Art digital museums were evaluated. These museums were classified as "content" digital museums because they held content from a number of sources. "Learning" digital museums were the third type examined, this included the Smithsonian Learning Lab, SHOW.ME, and the Web Gallery of art. These three digital museums presented content for learning and supported said content with learning aids.

The first digital museum examined was the Art Gallery of Ontario (AGO). This digital museum was the online representation for the Art Gallery of Ontario, located in Toronto, Ontario. The mission of the AGO was to connect people with art in order to facilitate new understandings about the world around them (Art Gallery of Ontario, 2017, Our Mandate). The main audience for this museum was tourists looking for information (English speaking only, the site information was not available in any other language), researchers, teachers, students, and families (resources were provided for these audience members both online and accessed at the physical location). The content presented in this museum included information about events, exhibitions, and collections found at the physical museum. When evaluating this museum with the heuristics there were some serious organization issues found. There were two main pages for this site, with two different labeling systems, organization systems, and search systems. The user moved back and forth between these two pages depending on the page they selected from a navigation menu.

The National Portrait Gallery (the physical location is located in London, UK) was the next digital museum to be examined. The national gallery was focused on improving their digital environment, put for by both the ten-year vision and digital strategy documents. For example, the National Portrait Gallery's digital strategy had three aims of improvement – Access, Understanding, and Sustainability (National Portrait Gallery, Nov. 2016, Digital Strategy). The audience for this digital museum included visitors to the physical location (including international users, as there were visitor guides in several languages), students, teachers, families and researchers. This site also endeavored to accommodate all users with accessible features (both digital and physical). For example, it had large print guides in some of the webpages. The content available on this website included images, information about exhibits and events, learning resources, etc. This site was very complex with multiple sections and cross-listed pages, and it had a varied amount of content organized within the polyhierarchical structure. See Appendix M for the National Portrait Gallery museum report.

The Cleveland Museum of Art (the physical location in Cleveland, Ohio) was the last brochure digital museum examined during the heuristic evaluation. This museum aimed to fulfill its role as both a leading art museum and one of Americas most distinguish northeastern cultural institute, while providing their audience with access to art according to the highest standards (Cleveland Museum of Art, 2017, Our Mission). The audience for this museum was focused on the physical museum visitors and researchers. There were a number of visitor guides available, in eleven different languages. There was also a rich collection of research and archived material on a subsite within the museums (for examples items that won blue ribbons in a 1930s fair could be searched for). The content on this digital museum includes visitor information for the physical museum, digitized collection items, information about events and exhibits, etc. This site had the most search systems (nine in total), and all of these needed to be examined during the heuristic evaluation.

The first of the three content museums evaluated was the Art UK. This digital museum began as a small charity, but now contained over 200,000 items from every public collection in the United Kingdom (Art UK, 2016, Welcome). The mission statement of Art UK described an organization that aimed to bring the collection of the UK to the public to enhance their knowledge, engagement, and enjoyment of art (Art UK, 2016, Our Mission). They aim to fulfill this mission through digitization, access to interesting content, using technology, crowdsourcing metadata (tags), and building partnerships (Art UK, 2016, Our Mission). The audience for this digital museum was very broad. The information digitized in the website can be useful for both art experts and novices alike. The licensing information available for images meant that additional audiences would be drawn to this site (those that wanted to use images – students, teachers, corporations, etc.) and the shop provided the option to purchase image licenses as well. The content in this site included artwork, stories about the art in the digital museum, events, and blog posts.

The second content museum evaluated was WikiArt. This digital museum's aim was to present and provide access to art as they tried to digitize the whole of art history – from cave artworks to modern private collections (WikiArt, 2017, About). They will reach this goal with the help of users; members of the WikiArt digital museum could add objects, tag them, and use tools to translate the information (WikiArt, 2017, About). This digital museum had a vast audience; they aimed to support anyone with an interest in art. This was supported by the number of languages that this site could be translated into (English, Dutch, Spanish, French, Portuguese, Russian, and Ukrainian). The site and content could be translated if someone translated the content already (if not they remained in English). WikiArt provided 30 language options to translate information into. The WikiArt digital museum contains information and images about artwork and artists. The site used metadata to divide both artists (by name, art movements, schools or groups, genres, nationalities, etc.) and artworks (styles, genre, media, etc.). There was more "content" available in the backend of the site. WikiArt let users edit existing records, add artwork, translate, etc.

The final content museum examined was Discover Islamic Art. It should be noted that this digital museum was created by the same organization (Museums with No Frontiers) as the digital museum evaluated in the 2012 pilot study (Discover Baroque Art). The aim of this organization was to present artifacts to users and have them act as ambassadors of the civilizations they represent (Discover Islamic Art, 2017, About Museums With No Frontiers). The audience of Discover Islamic Art was international. This was apparent by all the languages that this site (or parts of the site) could be translated into. The database could be translated into Arabic, English, French, and Spanish. The 18 virtual exhibits could be translated into Arabic, English, French, German, Italian, Portuguese, Spanish, Swedish and Turkish. Though it was important to note that the My Collections section (used when users signed in to the Discover Islamic Art) was only available in English. There was also a section set aside specifically for learning and schools. There were activities (like learning how to classify objects) that students could use to learn more about the objects in the virtual museum. These activities were definitely directed towards elementary school students. In this section there was also a page called "teachers zone" which produced a popup about information for teachers on how to utilize the teaching aids. This digital museum's content contains artifacts, images of architecture, exhibitions, introductions to Islamic artistic themes, and learning aids.

The first learning digital museum evaluated was the Smithsonian Learning Lab (part of the Smithsonian, but evaluated on its own). This digital museum brought together the learning collections from 19 museums, 9 major research Centers, and the National Zoo (Smithsonian Learning Lab, 2017, About the Smithsonian Learning Lab). The Smithsonian Learning Lab had vast networks of information resources and experts designed to bring excitement, discovery, and creativity to every lesson being taught (Smithsonian Learning Lab, 2017, About the Smithsonian Learning Lab). This site was aimed primarily at teachers; there was even an entire section in the "Help" pages dedicated for their support. This digital museum let teachers create their own learning collections and assign those collections to students. The target audiences were confirmed by looking at the different "Age Range" that the Learning Lab Collections were classified as. For example, there were categories like "Preschool (0 to 4 years old)," "High School (16 to 18 years old)," "Post-Secondary," and "Adults" (plus everything in between). So while this digital museum was aimed towards teachers and students it supports people of all ages who are interested in learning. The content in the Smithsonian Learning Lab was categorized into type (Image, Audio, Video, Text, and Learning Resources) and organized beneath the labels "Resources" and "Learning Labs."

SHOW.ME was the second learning digital museum evaluated. This digital museum brought together games, collections, places, and exhibitions from different museums, websites and archives across the United Kingdom. For example, the game "Art Lab" was part of the Tate Modern Kids site. The main audiences for this website were students/children (children and young people based on the content of the site) and teachers. Located at the top of the page was a graduation cap icon, for teacher resources. This brought up a list of "Stuff to read," organized for that specific audience. For example, the article "How to create illustrated books," which took them to a workshop that guided them through how to teach that project (aimed at grade 5 or 6). "SHOW.ME also featured editorial content written for children and young people" (Kennedy, 2014, Introduction for Teachers). This was seen by the websites that linked to the SHOW.ME site. For example, the "Make and do" article "How to make an Ancient Roman Snake Bracelet" came from the British Museum's website in the Young Explorers' section. SHOW.ME organized content by subject and type, including Art, History, Science, Dinosaurs, etc. The types of content found on the SHOW.ME website were collections, stuff to read, events, exhibitions, games, videos, websites, and places. These all fell within the subjects.

The final learning digital museum (and last museum in this study) was the Web Gallery of Art. This digital museum was "… a searchable database of European fine arts and architecture (8^{th}-19^{th} centuries), currently containing over 43,000 reproductions" (Web Gallery of Art, 2017a, Entrance Page). This was "… a free resource of art history primarily for students and teachers" (Web Gallery of Art, 2017b, Homepage). The digital museum supported learning and teaching with learning aids, found throughout the digital museum. For example, there are guided tours that exposed users to new ideas/themes ("Art in Spain in the 12th – 17th centuries"). Additional tools were useful for researchers, which included "Dual Mode," which let users view two images side by side for comparison (e.g. by subject matter, companion pieces, influences, etc.). While students and teachers were the target audience, the digital museum could also be "… a source of artistic enjoyment; a convenient alternative to visiting a distant museum, or an incentive to do just that" (Web Gallery of Art, 2017b, Homepage). As mentioned above, the content in this site included thousands of images, artist biographies, and learning aids (tours, dual mode, glossaries, etc.).

Contextual frameworks differ between websites, especially in the area of context (business/organizational information). However, the nine digital museums did have similar contextual frameworks within the defined audiences and content. The main audiences for all the digital museums were students, teachers, and researchers. These audiences were defined by the content on the site and often audience specific organization schemes, which explicitly stated those audiences (in the brochure and learning digital museums as well as Discover Islamic Art). In the content digital museums Art UK and WikiArt, these audiences were implied by the content and the fact that those digital museums wanted to reach the broadest audiences possible (thus encompassing those user groups). The National Portrait Gallery, Cleveland Museum of Art, WikiArt, and Discover Islamic Art also supported international audiences by providing content in multiple languages. Other audiences include tourists and families (for all the brochure museums). All the digital museums had examples of images, multimedia (videos, games, exhibits, events, etc.), and collections (the content). For example, online collections could be connected to a physical museum (the brochure museums), a country (Art UK), a theme (Discover Islamic Art and Web Gallery of Art), historical event (Smithsonian Learning Lab and SHOW.ME) and/or all of the above (WikiArt).

These digital museums were evaluated twice during data collection, first to become familiar with the digital museums and create the contextual framework (which informed the summaries above). The second examination occurred during the heuristic evaluation, examining the information architecture of digital museum with a list of heuristics (which were developed during the prior content analysis study). Once this study was completed, data analysis began.

Data Analysis

Data analysis of the heuristic evaluation research began by double-checking the findings reported during data collection. This meant that each digital museum was evaluated again, confirming the findings stated in the reports. Once the digital museum reports were confirmed, the collected heuristic evaluation data was examined in order to find similarities between the digital museums information architecture features. The comparisons examined similarities and differences between the types of digital museums (brochure, content, and learning). Allowing the data to be compared and contrasted, which was how the select digital museum criteria were created (discussed in "Findings" below). The bugs/issues found in the nine digital museums were also re-examined during the third evaluation (to ensure there were no false positives).

Brochure Museums

Organization Systems

All three of the brochure museums used hybrid organization schemes. The different organization schemes included both exact and ambiguous schemes. Ambiguous schemes were generally used as top and second level schemes (topical, task-oriented and audience-specific schemes). Exact organization schemes were more common in pages lower in the hierarchy, those that organize the content (alphabetical, chronological, geographical, and format). For example, the National Portrait Gallery (NPG), and the Cleveland Museum of Art (CMA) both used topical and task organization schemes for their top-level categories (found in the global navigation), while the Art Gallery of Ontario (AGO) used those two schemes, as well audience organization scheme. In the AGO this was represented by the category "Visitor Information," both NPG and CMA used the task organization scheme ("Visit" – the verb) to organize this information under. Exact organization schemes were represented in each museum. For example, all three of the museums used the chronological scheme to organize events and exhibits (by date and from the most current to future events). Content was organized in different ways, for example the CMA used topical, format, and geographical schemes to organize collections (e.g., Decorative Art and Design, Photography, and Korean art). Alphabetical organization was commonly found in A-Z indexes.

The organization structure for all three brochure digital museums used a hybrid structure. They all used a polyhierarchical structure with subsites scattered through the structure. These also had a database-oriented structure to organize the content for users. However, in the AGO, the database-oriented structure was only present in the Library and Archives and Buy tickets subsites. These sites were polyhierarchical because content was listed and accessed beneath different top-level categories. For example, in the NPG, the subsite "First World War Centenary" was organized both under "What's On" (as a second level category), "Group Visits," (third-level), and under "Collections" > "Explore Further" > "20th Century portraits" > "The Great War in Portraits" (4th level). In the CMA you can access the full "Exhibitions" hierarchy beneath "visit," "exhibitions," and "art." Subsites found within the three brochure museums included "Archives and Library" (NPG), "Museum Archives "(CMA), and "Library and Archives" (AGO), which linked to subsite database-oriented structures that accessed library catalogues. The database-oriented structures used the metadata of the content to structure and present the information for the user. The diagram below presents the organization structure and labels (the top three levels) for the Cleveland Museum of Art (to provide a sense of the organization and structure of the brochure museums) (See Figure 2).

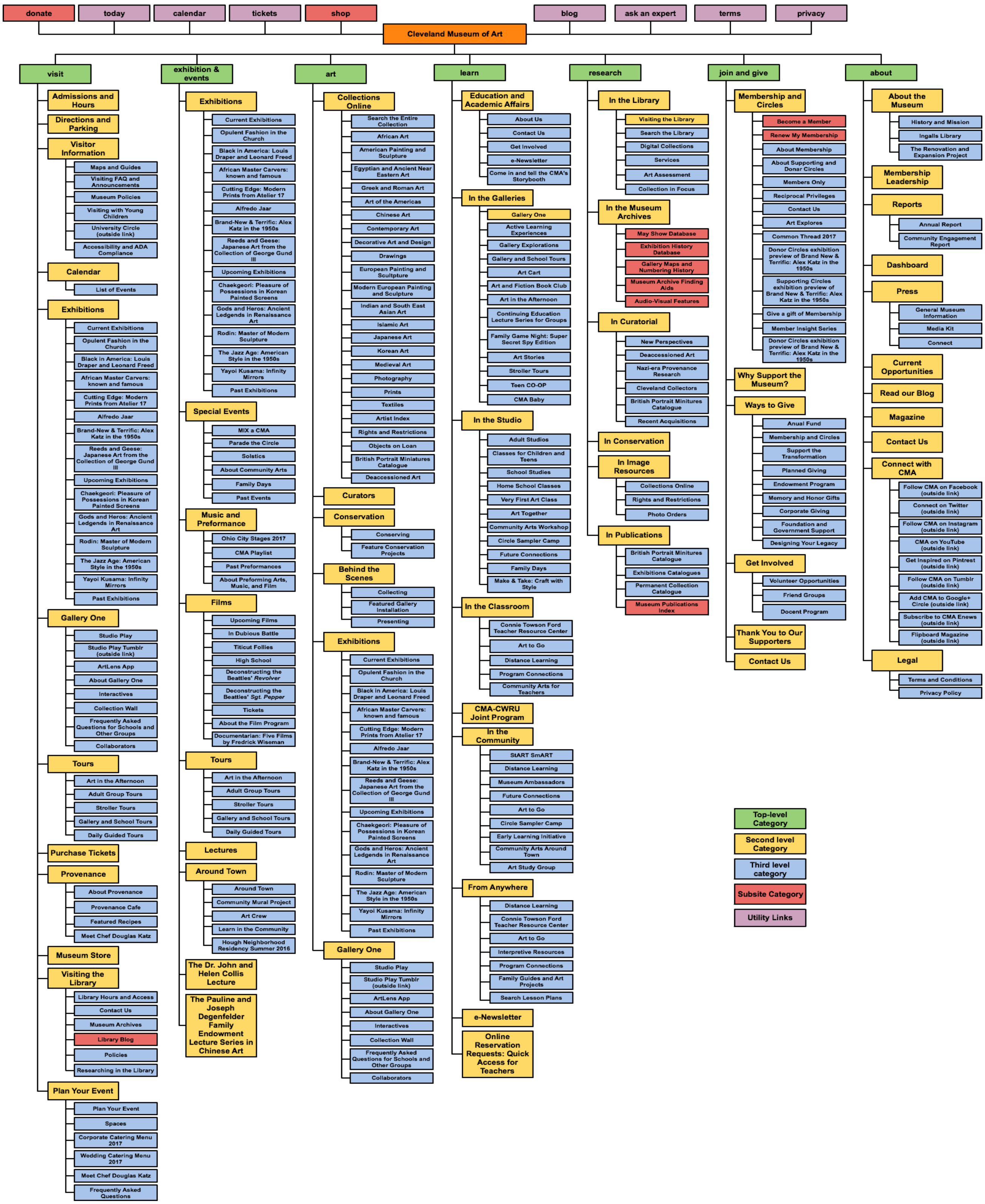


Figure 2. Diagram of Cleveland Museum of Art Digital Museum. The organization structure and labeling diagram for the Cleveland Museum of Art.

Labeling Systems

Each of the three brochure digital museums contained examples of every type of labels (contextual link labels, headings, navigation labels, index term labels, and icon labels). Contextual link labels were found organized under sections called “Related” (NPG), “Related Content” (CMA), and “You Might Also Like” (AGO). The links beneath those titles were contextual links, but the titles themselves were heading labels. Other contextual labels included inline links; the label was represented by sections of text. The inline contextual labels were distinguished by colour (represented by the sites graphic design), for example contextual links in AGO were red and in CMA and NPG they were blue. Heading labels were present as page titles, section dividers, and/or headings, which provided context to the information organized below. For example, in all three sites, the titles were the heading labels for that page. Navigation labels were present throughout all three sites. From the global navigation down to the contextual links, navigation labels were the most common labels in the three brochure digital museums. For example, in the NPG and CMA digital museums, their local navigation (in most cases) combined images with navigation labels. This provided context and introduced the audience to their collections (for example a painting of Anne Boleyn combined with the navigation label “Conservation of Anne Boleyn”). Index term labels (represented by tags, metadata, keywords, etc.) were commonly found within the content of the site (images and blog posts). For example, the blog posts in the AGO, NPG, and CMA all had tags associated with certain stories (e.g., Video, Education, Photography, etc.). Additionally, index labels were found in the metadata fields (discussed in depth below). Icon labels in these sites were used to represent links to social media (e.g., Facebook, Twitter, Instagram, etc.) as well as the envelope icon (email) and printer icon.

Each of the brochure museums had issues with labeling consistency. This ranged from grammatical errors to presentation problems (labels changed from page to page). The AGO had the biggest issues with their labeling consistency. The labels used at the top-level of the site changed as you moved through the AGO. You could see the difference between the two “types” of global navigation below (see Figure 3, the top image was the homepage, but the bottom navigation bar was where most of the destination pages were). The labels change between the two versions of the global navigation. For example, “Visitor Information” became “Visit,” or “Shop & Dine” became two different top-level categories “Shop” and “Dine” (when you select “Shop & Dine” you were sent to the “Dine” page in the second design). There were many more labeling issues found between those two navigation bars (the bottom global navigation bar was not a subsite, it seemed as if the website was in mid design change and it kept the old site while adding a flasher new homepage).

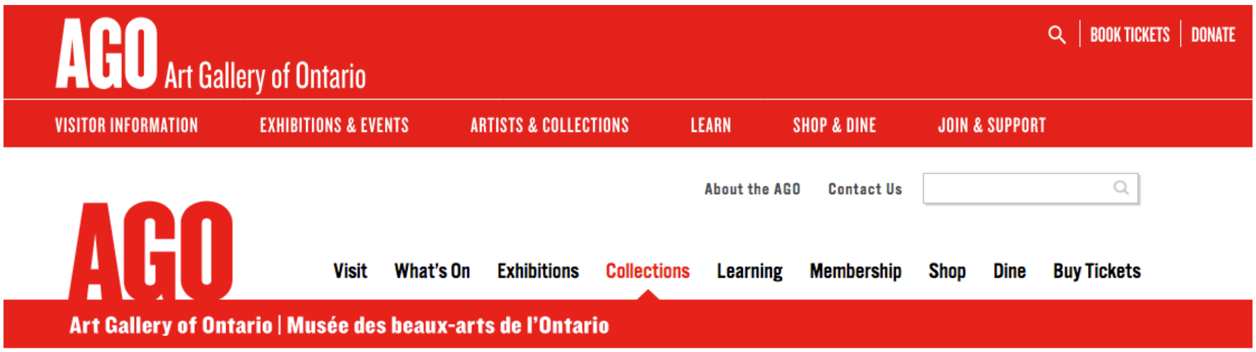


Figure 3. Global Navigation of the Art Gallery of Ontario. The top image was the homepage of the Art Gallery of Ontario. The bottom was the page where most of the Art Gallery of Ontario content was located.

The label consistency issues were less severe in the other brochure museums. The NPG had navigation labels that linked to pages that had completely different heading labels. For example, “Portraits on display” went to the “Room by room” page. This got even more complicated because “Portraits on display” within the “Tudor section” lead to a subsection within the “Room by room” page (inconsistent linkage). The labeling consistency issues in the CMA include capitalization problems (Top-level items were not capitalized; lower sub levels were) and terminology issues (links to the CMA blog were called “Read our Blog,” “blog,” and “CMA Playlist” respectively). This was one of the biggest issues found in these sites and definitely an area of information architecture that could be improved upon.

Navigation Systems

Each of the three brochure museums had all of the main types of navigation (global, local, contextual/hypertext, breadcrumb, and utility navigation). Each had a global navigation bar located at the top of the page (below a banner and digital museum logo). In general the global navigation for all three of these sites had a similar structure (and organization schemes, see above). For example, each of them had categories for visitor information (“Visit” or “Visitor Information”), event and exhibition information (“exhibition and events,” or “What’s On”), collections (“art,” “collections,” and/or “Artists & Collections”), education resources (“Learn,” “research,” and/or “Learning’), and membership information (“Join & Support,” “join and give,” and “Membership”). The NPG and AGO had links to the museum digital shop and the CMA has a link to the about page (all of the brochure museums had about and shop pages, they just did not all have them on the global navigation bar).

All the digital museums had local navigation, but the structure of these navigation features depended on the site (and in some cases the local navigation differed between different pages within the sites). In the AGO, local navigation was found for each subcategory along the left side of the page. These second-level categories were colour coded (e.g., the Thompson collection was blue and the Canadian collection was orange). In NPG, the local navigation was structured two ways (a) beneath the top-level categories there were navigation pages (a page of artfully arranged links, which combined portrait images, description text, and/or additional links) and (b) within the sub levels in the Learning section organized into three columns, populated with the sub categories (towards the top of the page below a secondary banner). The local navigation structure in the CMA was organized within a “homepage” for each of the main categories. This had a “Homepage” link first (an image of the CMA logo) with the rest of the links organized after (with a navigation label and an image).

The three brochure museums had contextual/hypertext navigation, which were organized into sections within a webpage (either at the bottom of the page or to one side) and were found as inline links. In the AGO, some pages organized contextual links in a box at the bottom of the main sub category page. For example, in the Spoliation Research page there were additional content pages (like the introduction to that section as well as interesting case studies of spoliation research) linked at the bottom of the page. Like the links organized at the bottom of the page, inline links were also highlighted red (reflecting the AGO design scheme). The NPG contextual links were organized under “Related” or “Further Links” and within the text. Additionally there were contextual sections at the bottom of select pages (with images, title, description and "Find out more" links). The CMA contextual and hypertext links were organized at the bottom of the page in individual bordered sections and there was inline links. Both of these were blue and did not change colour when selected.

The AGO, NPG, and CMA all had a version of breadcrumb navigation. In the AGO, beneath the global navigation bar, there was a "back" navigation link, so users could travel to the next level up in the hierarchy (while it was not a full version of the breadcrumb navigation, it was a similar example). The NPG breadcrumb navigation was located at the top of the content page (below the banner), though there were some issues with it. The polyhierarchical structure of NPG meant one could access subpages and content pages in multiple sections, however the breadcrumb navigation showed the path closest to a top-level category. For example, when selecting the "What's On" link, under visit, the breadcrumb navigation displays "Home/What's On" not "Home/Visit/What's On." In the CMA, the breadcrumb navigation was present on a few pages; this was the most straightforward version found in the three digital museums.

All the brochure digital museums had utility navigation (located in the in the footer of the website). The AGO had the same issue seen in the global navigation section; the labels and options changed depending on what page users were in (e.g., Visitor Information, became "How to get here" and "Hours and admission" links). The NPG utility navigation options included links to the NPG blog, social media, information about the NPG, Business and hire, National and international, and links to other languages (visitor guides). The CMA utility navigation was located in two places, at the bottom of the page, in a lower left side popup option, which contained links to purchase tickets, social media links, and information about the site (e.g., privacy policy and a sitemap). The other utility navigation options (along the bottom right side of the website footer) included donate (subsite), calendar, tickets, shop (subsite), ask an expert, term, and privacy.

These three digital museums differed in the type of supplemental navigation features they offered users. The AGO had A-Z indexes (located in the Malcolmson Collection for artists and in the Special Collection for collections), guides (when you purchased something in the ShopAGO and AGO tickets), control panels (used when joining the AGO, signing up for the AGO newsletter, etc.), and pagination navigation (when looking at content, on the main page to show site highlights, etc.). The NPG had multiple A-Z indexes (e.g., the hand list of names in the Reference collection, Photographic terms, Artist suppliers, Restorers, etc.), guides (quizzes "Shakespeare" and "Votes for Women," as well as "History of Hair and Beauty," which guided users through the beauty steps of different eras), control panels (when you joined up for membership), and pagination navigation (to scroll through highlighted elements, etc.). The CMA was the only brochure museum that had a site map though it did not display the full hierarchy (only the "visit" top-level category was expanded). It also was the only brochure museum that attached toolbar navigation to the images in the collection. These had additional actions that users could do with the images (post them on social media, email them, print and download). The CMA also had A-Z indexes (an A-Z list of artists), guides for purchasing tickets, control panels (for signing up for classes, events, etc.), and pagination navigation.

Each digital museum had variations of different advanced navigation; though none of them had personalization navigation and they all had visualization navigation. The AGO had customization navigation (located in the subsite, The Lodz Ghetto Photographs of Henryk Ross, the user needed to sign in in order to customize the site), visualization navigation (a map visualization and a keyword tag cloud in the Malcolmson collection), and social navigation (The Art Matters blog allowed users to use social navigation to browse the blog posts with tags). The NPG used visualization features for navigation (e.g., under the Tudor and Elizabethan subpage there was a navigational timeline), and social navigation (the NPG blog allowed users to navigate the content with pre-existing tags). Finally, the CMA used visualization navigation in two ways (a) tag clouds and (b) a geographical map that let user see where loaned collection items were in the world. The CMA also used social navigation (in a subsite users added tags to images, though you had to identify yourself to add them).

Search Systems

The AGO, NPG and CMA all had multiple search systems. The Cleveland Museum of Art had the most of any of the evaluated nine digital museums at nine separate search systems (for the collection, the website, the Library archives, databases, and the shop search). The algorithms for these search systems were unidentifiable (this was true for all the digital museums evaluated). The recall and precision varied depending on the museum and the search system evaluated. The search systems in the AGO generally prioritized recall over precision (the main/collection, ShopAGO, and Buy Tickets search systems). This was deduced, for example, when searching for “Group of Seven” 748 results were returned (some of the results only had the word “group” in them). However, in the Library and Archive search system precision was prioritized. The search for “Georgia O’Keeffe” returned only 98 results (in a large catalogue this was not very much) and every item was relevant to the search term (Georgia O’Keeffe was in the title or a subject). The NPG prioritized precision in both the main/catalogue and archive and library search systems. For example, in the archive and library search system, when you entered a search term you selected the results that you would like to see. If you searched for “Horse” you could then select the results you wanted to see with the associated keywords (e.g., “Horse in art,” “Horse breeds,” etc.). In the CMA seven of the search systems (collection, five databases, and shop search systems) emphasized precision. In the collection search, the query “Monet” returned results limited to works by Monet or his contemporaries/artists he influenced (and only 17 of the 73 results did not had Monet as the artist). The website and library archive search systems prioritized recall. In the website search, “Monet” returned many different results, but the last results were returned because they contained the word “monetary”).

The placement of the search interface depended on the search system. Though most of the search systems were accessed at the top of the page and (once a search was conducted) at the top of the search results (in all three digital museums). The library and archive catalogue search system interface was accessed on the main page, in the middle of the page (also in all three digital museums). The search interfaces (box and button) were all similar between these three digital museums as well. They had a search box with a magnify glass icon button. The only variation was that the CMA search box had rounded corners.

Advanced search was available in all three brochure digital museums. In the AGO an advanced search was available in the library and archive and shop search systems. The advanced options in the shop website include searching by keywords, departments, price range, etc. In the

NPG there was an advanced search system in the main collection and archive and library search systems. In the NPG archive and library search systems there were multiple advanced filters (e.g., format, location, title, etc.). In the CMA digital museum, the library and archives (as well as a few of the databases) had an advanced search system. For example, in the library and archive search, you could filter results by keywords, publication, number fields, etc. Advanced search features were more commonly found in catalogues that had content with rich metadata and many items.

How the search results were displayed depended on the digital museum, but most of them presented results in list form (with ten items per page). Variations on this presentation included that, the NPG allowed users to view results 20, 40, or 60 items per page and the CMA collection search system did not use the list format, but presented returned objects in a thumbnail grid (all images). The information displayed for these results encompassed a title, description, artist, images, etc. The display reflects the metadata that the content contains. For example, the search results in the CMA Shop displayed an image, title, description, and price with the search term highlighted.

The only type of ranking search results used was relevance. For example, in the AGO the main/collection search system uses relevance to rank the results, the first result in a search for "Group of Seven" was the Canadian Collection and an old event specifically about the Group of Seven. Many of the results were out of date (e.g. an event from 2005), but included in the results because they were relevant to the query. Other search systems that ranked results by relevance included the CMA collection and the NPG main collection searches. There were many different types of sorting methods, but the ones used most commonly in these three digital museums were alphabetically (by title and artist) and chronologically (by publication date, etc.). For example, the AGO library and archive search systems let users sort search results by none, artists (alphabetically), media, publication date (chronologically), and title (alphabetically). Often these sorting methods added the option to sort in ascending or descending order (e.g., A-Z or Z-A). The search systems that did not have any sorting options were the NPG and the CMA shop subsites.

There were many additional options included in the three brochure digital museums. The most common additional action in all the search systems evaluated was the ability to conduct a new search easily (done by placing a search interface at the top of the results). Often this included retaining the search term queried, making it easier to do a new search. The AGO provided additional actions in the library and archives and shop AGO search systems. For example, the library and archive search system let users save/bookmark the returned items to a list (one could do this by individual items or by bulk – entering a number range, up to 100), as well as saving the search by emailing it to yourself (and you could choose the format the content was saved in – HTML, Plain text, delimited, MLA and Chicago). The NPG main search system let users refine their search (using a drop-down menu) and the archive and library search system additional options included saving a subset of the results to a "user list," save the search results, print and email the search results, and narrow the results down (using related terms). The additional actions available in the CMA were that users could save individual search results (the collection and museum archive search system) and narrowed search results using filters (the collection and library archive search system).

The query builders available in the search systems varied between systems and between the brochure digital museums. However, within the brochure museum there were no examples found of phonetic tools or natural language processing tools. There were only two search systems that included spell checkers, the main page search system in the AGO and the library archive search system in CMA. For example, in the NPG collection search system, the "Earl of sandwhich" did not return any results, but the "earl of sandwich" returns 63 portraits and 7 people results.

The search systems often had some stemming capabilities, including the CMA library archive search system (e.g., a search for "painting" returns "painters"), the NPG archive and library search system (e.g., a search for "paintings" returned results with "paint"), the main search in the AGO (e.g., a search for "impressionism" returned "impressionist"), the AGO library and archive system, the ShopAGO system (e.g., a search for "Canada" returned "Canadian" and "Canadianism"). The only forms of autocomplete/autosuggest available were past search queries, accessed in a drop-down menu when users started to type in the search box.

For example, in the CMA website search system it will suggest past searches (even if they were spelled incorrectly).

Each of the three brochure museums provided advanced query language in at least one of the search systems. This included Boolean language and stripping out the stop words (a, the, etc.). For example the NPG main/collection search system supports Boolean operators, but it did not strip out stop words. The ShopAGO did strip out stop words though. Often Boolean operators were available in advanced search interfaces (they were used to separate search fields). For example, the CMA library archive search system used Boolean operators as well as wild card character (*) to support stemming.

The content most often indexed for searching was the metadata of the objects (titles, description, keywords, etc.). The returned results were all destination pages, the actual content (instead of navigation pages). Filters represent the search zones (either in the advanced search function or on the side of the page once the search had been conducted). For example, in the CMA library archive search system the advanced search allowed the user to define the sections they would like to find their search query within. This included subsections beneath keyword, publication, auction related fields, etc. Other indexing methods included full-text indexing (found in the AGO main search – when blog posts were returned), indexing for recent content (this method of indexing was found in the NPG events search system), and indexing by topic (the metadata indexing). Vertical searching was used in the NPG main/collection search system. When users entered a search query (e.g. "Equestrian"), the user could then choose to view the results found in particular section of the website (this included "People," "Portraits," "Events," and "Other Pages"). Other search systems with this functionality were the AGO main search and the CMA website search system.

Vocabulary Systems

All the digital museums evaluated had metadata, though the amount and detail differed from site to site. The types of metadata were descriptive, structural, administrative and embedded (often a combination of the first three, located within the websites HTML code). Due to the vast amount of content held by the three brochure museums there were different metadata fields for different types of content, these included images, events, exhibitions, blog posts, etc. Descriptive metadata organized information about the content like the title, description, keywords, subject and themes (unique to the NPG), and much more. The most common descriptive metadata found in the three digital museums were title, description and keywords. However, the AGO collection content only had a title and the CMA had the additional field of disciplines in learning lesson plans.

The administrative metadata (information about how the items were managed (Wodtke & Govella, 2009)) included items like artist/creator, data of creation, date of publication, price, location, time, current owners and item/catalogue number. For example, the administrative metadata for the portraits in the NPG digital museum had the artist, date, NPG call number, and provenance metadata. The images in the CMA included date, artist/creator, IC/call number, credit line, provenance, and exhibition history. Finally the administrative metadata found in the AGO collection items had creator attribution, date of creation and the current owner. Administrative metadata varied the most, depending on the digital museum being evaluated (business organization effected this metadata).

The structural metadata depended on the content and the file size. This included size (file or content), medium/material, and type/format. For example, the objects in the NPG shop had the structural metadata of medium (materials) and size (though this depended on the item being viewed). The AGO collection items had the size and medium for the works (this also varied between types of content). Finally, the CMA structural metadata for items in the shop had size and format.

There was not that much embedded metadata in these three digital museums. The AGO used embedded metadata (to facilitate searching) in ShopAGO. For example, the keywords connected to a “Haring White Cycle T-Shirt” included “Large,” “Artists,” Modern,” “TSH,” “POP,” etc. (though they did not all make sense and some of them were duplicated). The items in these three digital museums also had open graph metadata, which turned the item into rich graph objects, so that they could then be shared on Facebook pages (The Open Graph protocol, 2014).

The open graph metadata found in the HTML included og:sitename; og:type; og:url; og:title, og:description; og:update_time, and og:image. The Cleveland Museum of Art also had Dublin core embedded metadata. The types were dcterms.title; dcterms.creator; dcterms.description; dcterms.type; dcterms.format; dcterms.identifier; dcterms.language. Embedded metadata helps improve the indexing and discovery of items in the digital museum.

The controlled vocabularies of the digital museums were the hardest information architecture principles to detect and to determine. It was important to note that this section involved the interpretation of the terms, labeling, and organization found in the site. Authority files were determined by any A-Z lists found on the site. Classification schemes were evaluated by looking at the facets found within searching and browsing systems. Finally, the thesaurus was determined by the keywords and glossary (if available) and then compared to standards and/or reconstructed (to show possible term relationships found within the used keywords). None of the brochure museums had synonym rings (that could be detected).

The AGO controlled vocabulary was difficult to detect. It did not seem to have a discernable controlled vocabulary. It did not appear that there were preferred terms for the content or the artists (for the main site). For example the spelling for Van Gogh varied (the “V” was capitalized in one and lowercase in the other). This was mainly due to content authors, but there did not seem to be a site wide preferred term for artists. Meaning that this digital museum most likely used an indexing thesaurus (created by authors), though the embedded metadata in the AGO was not robust enough to determine this.

The authority files in the three digital museums were encountered in the A-Z artist index that the sites had (and all three sites had an example of this, even though they may only be encountered in a subsite). The AGO Malcolmson collection had an A-Z list of all the artists; the preferred structure for those names was “Last name, First name. ” For example, “Bedford, Francis” or “Marey, Etienne Jules.” A site wide example was found in NPG, all artists and sitters (people who sat for portraits) were found within A-Z lists, so that users could browse content by them (and use them as references when searching). For example, Artists: “Sir Anthony van Dyck (1599-1641). 1018 Portraits” (listed under “V”) and Sitters: “Sir Frederick Augustus Abel, 1st Bt

(1827-1902), Chemist and explosives expert. 8 Portraits." The CMA had an in house authority file for artists – the Artist Index. Though the format of the names listed in the index did not always match how they were used in the metadata (e.g. "Caravaggio" was the listed name for that artist, but in the metadata it was "(Caravaggio, Italian, 1571-1610)".

The classification schemes found in the digital museums could be extrapolated from the facets or filters used to refine searching/browsing. These were arrangement of preferred terms that the site used to define content. For example, beneath the classification "format" there were preferred terms like painting, sculpture, photograph, print, drawing, textile, ceramic, etc. AGO had classification schemes in the events page and in the Boxwood Collection subsite. For example, the Events page allowed users to narrow events down by type (AGO Collections, Art Parties, Exhibition, Family Events, Food & Drink, Performance, Screenings, Shopping Events, Special Events, and Talks). It was important to note that those classification schemes were not site wide or available for all content. The classification schemes available in the NPG could be found in the advanced search (filters that users could use to narrow their search query). This included place (the options were Africa, Americas, Antarctica, Arctic, Asia-Pacific, Europe, and Middle East), professional category (Agriculture and Food, Art, Law and Crime, Medicine, Religion and Belief, Social Welfare and Reform, etc.), and medium type (drawing, miniature, mixed and new media, painting, photograph, print, and sculpture). These filters were site wide and applied to all content. The CMA classification schemes included collections (photography, prints, European paintings and sculpture, Medieval Art, African Art, etc.), by creator (authority file – list of artists drawn from the A-Z index), and type of object (photograph, sculpture, textile, painting, metalwork, etc.).

All the library and archive catalogues followed some types of thesaurus standards (these were subsites within the digital museums). The AGO library and archive catalogue followed the Library of Congress Authority File and Subject Headings (for the artists and subject). The NPG archive and library catalogue used the Art & Architecture (Getty) Authority file (e.g., "Caravaggio, Michelangelo Merisi da (Italian painter, 1571-1610)") and the Library of Congress Subject Headings (e.g., "Cardinals in art" or "Horse breeds"). The CMA library archive used the Library of Congress Subject Headings as well. The content within the websites did not follow any standard. They were compared to the thesaurus standards LCC, Art & Architecture, the Humanities and Social Science Electronic Thesaurus (HASSET), and the artist/creator names in VIAF.

The thesaurus in the AGO was indecipherable (mostly due to the lack of keywords in the metadata of the objects, embedded or otherwise). The NPG organized the terms that described portraits in a subpage called "Subject and themes," that the thesaurus of terms used for searching, browsing and indexing could be extrapolated. For example:

- Genre
  - Children
    - Royal Babies
  - Couples
    - Double portraits
  - Family
    - Family portraits RT Group portraits
  - Group portraits
  - Nudes and naked figures
    - Body
  - Self-portrait
  - Weddings
    - Wedding inspiration

From this possible thesaurus structure the semantic relationships between the terms could be viewed. For example, the hierarchy relationship was viewable in this passage "Genre > Couples > Royal Babies." Related terms (associative relationships) had also been explained in this list as well. The semantic relationships between the thesaurus terms were informed by the thesaurus standards used.

All three of the digital museums had examples of faceted classification (as mentioned above). The AGO allowed user to narrow down the events they would like to see by type (GO Collections, Art Parties, Exhibition, etc.) and additional filters ("Free Events," "Georgia O'Keeffe," "Member Exclusive," and "Events after 5 PM"). The NPG blog had a type of faceted classification. Users could limit the blog posts they wanted to see according to the tags that had been applied to them. The available facets included behind the scenes (conservation, discoveries, exhibitions, etc.), Portrait Period (Tudor, Stuart and Civil War, etc.), and Medium (painting, sculpture, etc.). Faceted classification could also be found in the NPG advanced search; these filters (place, profession, medium type, etc.) could be used to narrow the search results down with. The CMA allowed users to use faceted classification to narrow down search results (in the collection and website search system). This included (as discussed above) collections (photography, prints, European paintings and sculpture, Medieval Art, African Art, etc.), by creator (authority file – list of artists drawn from the A-Z index), and type of object (photograph, sculpture, textile, painting, metalwork, etc.). Additional facets the user could select were “Object on view,” “Object with images,” and “Museum Highlights.” Faceted classification enriched the user experience and helped them find information they were looking for.

Content Digital Museums

Organization Systems

The content digital museum all used hybrid organization schemes. The digital museum’s top and second levels used topical and task based organization schemes. For example, Art UK had “About” (topical) and “Discover” (task-oriented). WikiArt top-level categories included “Actions” (task) and “Artworks” (topical). The Discover Islamic Art digital museum included audience (My Collection) and an exact scheme, chronologically (Timeline). Additional second level organization schemes were alphabetical (WikiArt), format (Art UK and WikiArt), task (WikiArt), and geographical (WikiArt). The Discover Islamic Art permanent collection could be browsed by Country (geographical scheme), Period/Dynasty (chronological), and by start and end date (chronological). In Art UK, the Venues (institutions that the content had been digitized from) were broken down by location (first by larger areas, e.g., London, Scotland, etc. and then by counties and/or hamlets – geographical organization scheme). In WikiArt the subcategories organized information in two ways, by name (alphabetical) and by count (the amount of artists/artworks that were classified into each category). The second level categories that used this method included: Art movements, Schools and groups, Genres, Fields, Nationalities, Art Institutes, Styles, Genres, etc.

The content digital museums all contain hybrid organization structures. All of these digital museums had a polyhierarchical top-level structure with database-oriented bottom level structure, used to organize the content of the digital museums. The polyhierarchical structure found in Discover Islamic Art, Art UK and WikiArt used focus entry points to access content under several categories and in several locations. For example, the focused entry points in Art UK allowed users to access content by artist, artwork, stories and topics. In WikiArt, the focused entry points were within the subcategories "Artists" and "Artworks." This included Art movements, genres, popular artwork, media, etc. Another organization structure found in the WikiArt digital museum was hypertext structure (commonly found in wikis). In WikiArt users could access featured blog articles, artists, stories, events, etc. within content pages using contextual links. The database-oriented structure of all three sites used the metadata of the content to organize and present objects for users. For example, in Art UK all the works have a few basic metadata fields filled in – title, artist, date, medium, measurements, acquisition method, and accession number. Finally two of these sites had subsites (Art UK and Discover Islamic Art). For example, when users selected Learn with MWNF and My Collection in Discover Islamic Art they were taken to subsites that connected them with their account information or learning aids.

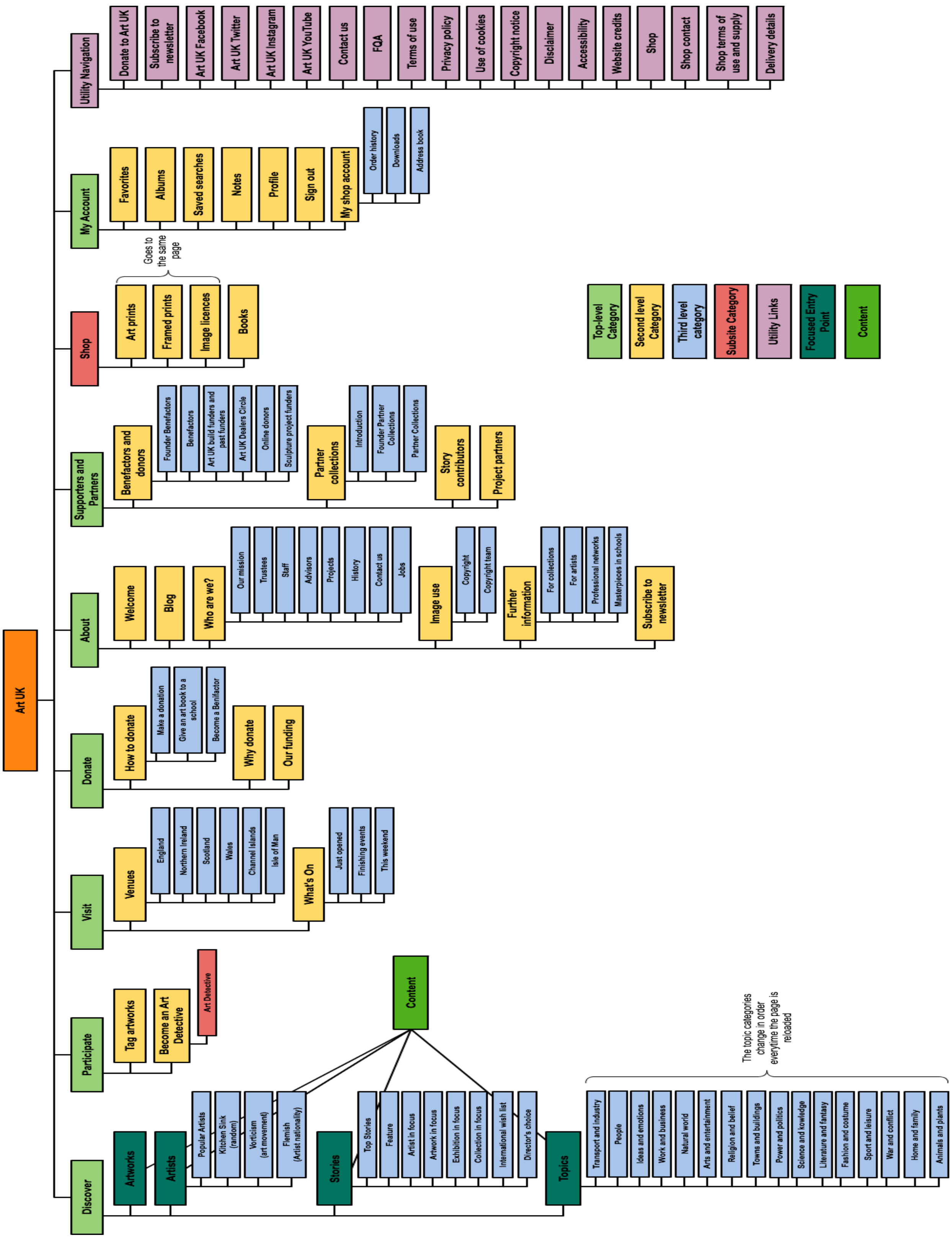


Figure 4. Diagram Of Art UK digital museum. This shows the pages and subpages organized within the Art UK digital museum, as well as the labels of the site.

Labeling Systems

The three content digital museums evaluated had examples of all the different types of labels (contextual link labels, headings, navigation labels, index terms, and icon labels). In the three content museums there were two types of contextual link labels, those were organized beneath heading labels that pointed users to more information about or related to the content they were looking at. In Art UK, much of the content was contextually organized, for example in the "William Hogarth" artist page the contextual links included other images by this artist, the stories that included this artist, the venues that contained works by this artist and outside links to bibliography information (Wikipedia, Oxford Index, Oxford art online (locked), and Oxford National bibliography). In the WikiArt, the contextual link labels included text that connected different artists together ("Related Artists"), different artwork ("Related Artwork") and artwork that was done by the same artist ("Famous Works"). These three content museums also had inline contextual link labels. Chunks of text represented the labels, distinguished by colour and/or underlines. For example, in Discover Islamic Art there were inline contextual link labels in the text pages, distinguished by a white underline (the labels themselves were part of the text).

The content digital museums evaluated used heading labels throughout their sites (to contextualize the information displayed below them). Examples of headings found in these digital museums included titles, section dividers (for text and objects), and headings for groups of links (like the "Related Links" heading found above). For example, in Art UK the heading labels appeared in the center of the top page banner. Those heading labels showed the second and third level categories in the hierarchy. Discover Islamic Art contained heading labels in the "Exhibition" and "Artistic Introduction" pages (the exhibitions were presented to users like a document or PDF – e.g., users navigated to the "next page"). Heading labels displayed the title of the exhibit and the section within the exhibit they were viewing. Heading labels informed users about what was organized beneath them (either text, content, or links) this included the art movements that were organized under headers in WikiArt. Users could not click on the headers, but the content beneath those labels were links that belonged within that art movement. For example, beneath Korean Art were the categories Joseon Dynasty, Korean Informel, and Dansaekhwa (Korean Monochrome Painting).

Navigation labels were the most common labels found in all three of the content digital museums (these included contextual link labels and index terms). For example, the global navigation labels in WikiArt were actually headers (they did not link to new pages, just provided context to the navigation labels organized beneath) with mega drop down menus that contained the navigation labels – they linked to content navigation pages. By choosing an option in the menu or selecting "view all" the user navigated to another navigation page. These were obviously found either in or below the global, local, and utility navigation features (see below for more information). The organization diagram above (Figure 4) showed the navigation labels for the Art UK digital museum. Discover Islamic art had navigation labels in the global navigation menu and organized on the right side of the page when viewing an object.

Index term labels (often represented by keywords, tags, etc.) were commonly found within the three content digital museums. In Discover Islamic Art, index terms were found in the glossary for the site (which defined the correct terms to use when searching/browsing). For example, under "G" there were seven entries (including styles, mediums, a location, etc.). In Art UK some of the index terms were organized beneath "Topics," so that the user could search using the predetermined keywords. Users of Art UK could also use tags added through crowdsourcing, which are index terms as well (created by the website audience). In WikiArt the genre index terms (for example) were listed together (organized by count and by name). When users clicked on them they were brought to a list of artists who created works described with the selected index term label.

Examples of icon labels were found in the three content digital museums. Generally these represented social media links (Facebook, Twitter, Instagram, Google+, etc.), but there were other icons as well. The Art UK used icon labels to imply actions user could do with the images in the collection (a heart added an image to a list of "favorites" and a file folder added an image to a collection) and the social media links (Facebook, Twitter, Instagram, and YouTube). WikiArt used icon labeling in the "My account" page. For example, within an album they were icons labels that let users edit the album (a pencil icon) or post it to social media (an arrow). Additional icon labels included those found in Discover Islamic Art, within the "Learning with MWNF" page (for the individual games or learning aids found there). For example, in the "Learning to relate form and function" game there were two icon labels associated with the images in the game – a question mark in a square (which linked to the images content page) and a magnify glass with a plus sign (used to open a bigger image of the item).

Each of the three content museums had labeling consistency issues (though these were minor issues compared to those found in the brochure museums). For example, in the Art UK local navigation bar (not global, it was only on first page) the organization of the categories goes Artists, Artwork, Visit, and About. However, the global menu lists Artwork and Artists under discover (in that order) and Visit and About were separate categories entirely. The order of the links differs depending on the navigation menu. The global main menu changed the label/navigation order when users were in the Shop (shop, generally located at the bottom of the menu was now moved to the top). There were also inconsistencies when using question marks (e.g., "Who are we?" vs. "How to donate," "Why donate," and "What's on"). Capitalization was the only consistency issue found in the WikiArt digital museum. Three of the categories did not capitalize the first letter of any of the index terms – fields, genres (under both artwork and artists), and media. However, the index terms in schools and groups, art movements, styles, nationalities, etc. all have the first letter of every word capitalized. This was awkward, but not detrimental to the site. The Discover Islamic Art digital museum had very few labeling inconsistencies. The only issue was that two of the global navigation labels did not describe the content that belonged beneath it well. For example, the Database top-level category was where users found the search feature in this site and under permanent collection was where users found the browsing feature. Initially it was thought that the search function would be found under Permanent Collection and did not know what would be found under Database. It should be noted that this depended solely on subjective interpretation and may not be an issue for other people.

Navigation Systems

Two of the content digital museums had all of the main navigation features (WikiArt and Discover Islamic Art). The third (Art UK) did not have the breadcrumb navigation feature. Global navigation was present in all three of the selected digital museums. In Art UK, it was located under the menu button (which had both an icon and text label), at the top right corner of the website. Once the user clicked on the button, the menu expanded left to reveal the top-level and secondary level categories of the site. The global navigation in WikiArt was a combination of heading labels and mega drop-down menus. For example, when users clicked on the "Artist" heading they were shown a mega-menu, which displayed the second and third levels. These were part of the global navigation because they were available on every page in the site. In Discover Islamic Art the global navigation bar was located at the top of the page above the banner (the most traditional version of the global navigation). The global navigation categories were "Permanent Collection," "Database," "Exhibitions," "Artistic Introduction," "Partners," "Timeline," "Learn with MWNF," and "My Collection."

The content digital museums all had examples of local navigation. The Art UK local navigation was found either within the category pages (images/links which were organized under headings) or beneath the banner window. This second type of local navigation was found in the business side of things (About and Donate). WikiArt had two local navigation features. The first local navigation menu was located within the Actions category. This option only appeared when the user was signed in (and in a content page). The actions included edit (edit an existing items metadata), add artwork (add a new item), translate (translate the items information – 30 available options), etc. The second local navigation method was when users entered the subcategory "view all." For example, in the media page (beneath the artworks category) there was a complete list of all index terms attached to different works of art. In Discover Islamic Art the local navigation was found within the "Exhibition," "Artistic Introduction," "Learn with MWNF," and "My Collection" categories. For example, within "Exhibitions," the local navigation listed all the exhibitions users could view.

Contextual/hypertext navigation was found in all three of the content digital museums. In WikiArt the contextual links were organized beneath the individual content pages. These included links that connected different artists together ("Related Artists"), different artwork

("Related Artwork") and artwork that was created by the same artist ("Famous Works"). The metadata in this site contained contextual links that connected an artist/artwork with other categories (e.g., Nationality, Art movement, Influenced by, Influenced on, Teachers, Wikipedia, etc.). In Art UK there were contextual/hypertext navigation links under stories (on the side of the page), a "Did you know?" section that provided links to additional information (e.g., a link to Art and Architecture Bruegel family tree in an story about his work). Within the Discover Islamic Art content pages were different types of contextual link labels, organized beneath "Related Content." The black arrow before it identified the contextual link. The contextual links were organized beneath heading labels (e.g., Related objects, On display in, See also, etc.). All the content digital museums had inline hypertext links, distinguished by colour or underlines. For example, Art UK had blue underlined links in the blog and story content.

Two of the museums had breadcrumb navigation (WikiArt and Discover Islamic Art). In WikiArt the breadcrumb navigation was only within the "My account" section. For example, in the favorites album the breadcrumb navigation feature would display "Home > user name > Album > Favorites." In Discover Islamic Art the breadcrumb navigation was in the "My Collection" sections. For example, if you decided to change your email address, the breadcrumb navigation looked like "My Collection > My Details > Change Email Address."

All three of the content digital museums had utility navigation, located in their expanded footer. Art UK's utility navigation had nineteen different links, including both in site links and outside links (Twitter, Facebook, etc.). This included a prominent link (coloured red) to allow users to donate to the Art UK digital museum. In Discover Islamic Art the utility navigation options included "About MWNF," "Contact," "Feedback," "Legal Notice," "Credits," "Language Policy," and "Cookies." The rest of the text in the footer contained copyright information. The utility navigation in WikiArt had links to "Home," "About," "Feedback," "Donate," "Terms of use," "Android App," and "Language."

None of the content digital museums featured all the supplemental navigation options detailed in the heuristics, but they all had at least two examples. WikiArt had multiple A-Z indexes (for styles, genres, artist names, etc.), guides to help users upload multiple images into the site, control panels for editing user and artwork information, and pagination navigation in the "Quick edit artwork" page. For example, in that section there was multiple control panels for each of the objects users could edit, pagination appear at the top and the bottom of the page to help scroll through the objects. The Discover Islamic Art supplemental navigation features included a glossary (A-Z index), control panel navigation (used to register an account in the Museum With No Frontiers), and pagination navigation (found in the search results, permanent collection, and timeline). Finally, the Art UK digital museum contained control panels (e.g., users could fill out one page forms to make a donation, contact Art UK, and change their profile information), and toolbars (allowing users to post the image they were looking at to Facebook, Twitter, add it to their favorites, etc.).

All the digital museums had advanced navigation features (though none of them contained personalization). The Art UK had customization, visualization, and social navigation. For example, Art UK allowed users to customize their account; they could favorite items so that they were easy to access and add images to their albums. Art UK's visualization navigation was available both on the main page and in the venue section. The Art UK used maps to locate artwork within participating venues. Social navigation in Art UK referred to how users could tag the images with their own keywords (after signing up). Though this feature was being updated during this study (and so could not be examined) users could still navigate by tags. Discover Islamic Art had customization and visualization advanced navigation features. Users could create their own collections by selecting and adding content (from either the database page or from within "My Collections"), customizing their accounts. Visualization navigation only appeared within the database section. When the user's cursor hovered over an item, a larger image appeared on the left side of the page (with additional metadata). Finally, WikiArt had customization and social navigation features. Much like Art UK, users could favorite items so that they were easy to access within their account, and they could add images to albums in order to group artwork. The WikiArt search system used social navigation methods, because users added the bulk of the metadata (the keywords and tags).

Search Systems

The content digital museums had less search systems then the brochure digital museums. The Discover Islamic Art digital museum had two search systems (in the main website and in the user account page), while Art UK and WikiArt each only had one. The search system in WikiArt weighed precision over recall. Every search result had the query in its title or in the artist field (some part of the name). Additionally, because there was a rich autocomplete/autosuggest query builder, the user could select pre-existing search queries, improving precision. Both search systems in Discover Islamic Art prioritized recall over precision. For example, a search for mosaic returned results that included buildings without the search query (in the metadata). This may be because the site did not have that much content (as compared to the other digital museums studied). Art UK emphasized precision over recall. In a search for "Rubens" there were 286 results and "Rubens" featured in every item, either as the artist or as an influence of (ordered by relevance). Precision was aided by predetermined search terms, which the user could select from a drop-down menu (when they started to type).

While there was not that many search systems, the content digital museums search interfaces were complex (and in one case the interface differed from page to page). For example, in Art UK the search interface was found in two locations (in every page, but the About section). The first was in the global navigation menu. This was a simple search box with a magnify glass icon button. The second was located beneath the banner at the top of the content pages (in the artwork, artist, homepage, venue, what's on, and shop pages). These had a search box with a magnify glass icon, along with additional features (the homepage search did not have search filters). The filters in the artwork search included region, type, topic, style, license, and image (if it had one or not). The filters in the artist search included nationality and popularity (a check mark – yes or no). The venue filters were region, type, and if it was open to the public (a check mark – yes or no). The what's on (events) filters were region, when, audience, type and if it was free or not (a check mark – yes or no). The shop search filters were collection, region, and type. The Discover Islamic Art search interfaces were both advanced search interfaces and contained a similar structure. It had three search boxes, separated with Boolean operators (just and/or) drop-down menus. Users could narrow the fields the search term was in (Keywords, Name, Locations, Provenance, Period/Dynasty, Patron, Architect/Artist/Master, Material/Technique and Other). Once finished crafting their query, users clicked on a yellow "Go" button. The search interface in

WikiArt was persistently located at the top of the page. The search box extended between the "Artworks" category and the "Action"/"My Account" options (depending on what page the user was in and/or if they were logged in). The search button was represented by a magnify glass icon.

Only Art UK and Discover Islamic Art had advanced search, though this was not represented by separate interfaces (it was built in to the main search interface). The advanced search functionality in Art UK was found in the filters that users could add to a search term (see above) and they could be used to combine search terms. For example, a search for "Titian" could have the additional search term "Andromeda," which would reduce the number of artworks returned from 139 to 2. Users could create these complex searches by adding more search terms to narrow the results as well as using the filters. Discover Islamic Art used an advanced search interface for both of its search systems (as the main search interface, see above).

How the search results were displayed depended on what content was searched and which search interface was used. In WikiArt, the results displayed depended on what was being searched. For example, if the query was selected from the autocomplete/autosuggest drop-down menu (and applied to both artists and artworks like a subject), then the results were displayed with pictures of the work of art (or artists portrait), title, artist, and date. Searching for an artist name in WikiArt returned different artists (with their name, picture, and date) at the top, and then artworks associated with the artist name searched (e.g. Rubens returns four artist entries and 438 artworks). Discover Islamic Art had two search systems, and they presented results differently. The main site search system displayed the results in a thumbnail/list form (two across, 12 per page). A large image with a little bit more metadata (title, date, location, and museum location) appeared on the left side of the page when the user hovered their cursor over one of the smaller images. The information that appeared with the smaller images was title, date and location. The second search system (in the user account subsite) displayed search results in bordered fields. There were 10 items per page in a list. The displayed information included an image, title, date, museum, medium, and location. The search button was a simple "Search" button. Art UK search results were displayed in a list view (though it was more similar to what other digital museums called thumbnail – images with title, date, and artist metadata) or in a map view (organized by the venue it was located in). The list view was the default, even in the venue search. The information displayed depended on which section of the search users were in – the Shop separated the results into artwork prints that could be bought and products (books), the artists biography page link first (if the user selected that option). An artist search returned the artists name, and how many pieces of art they uploaded on to Art UK. An artwork search returned the title, artist and date (most of the time). And the general search (within the global navigation menu) returned results from all categories – artworks, artists, stories, events, etc.

Discover Islamic Art and WikiArt ranked the search results by relevance (there were no other options to rank/sort within those content digital museum). For example, in WikiArt the top results all had the search term in the title (e.g., "Nymph" returned *Flower Nymph* as one of the first results). The search results could be sorted in Art UK. This site's search results were initially ranked by relevance (the default option), but they could also be sorted by date made: new to old, Date made: old to new, Title: A to Z, Title: Z to A, Artist: A to Z, and Artist: Z to A.

There were many additional actions available in the content digital museum websites (a by product of customization navigation). Discover Islamic Art additional actions included saving a subset of results (by adding them to a collection). Users could also narrow the search results down by choosing "Refine Your Search" and use the search interface to add additional terms (box, subject drop-down menu and a search button). There was also a "New Search" button, which brought users back to the advanced search page. In WikiArt users could save a subset of results by favoriting them or adding them to an album. Additionally, users could conduct a new search as the interface was easily accessed at the top of the page. Finally Art UK had many additional actions available, though most were reliant on being an account member. Users could save individual search results into a favorites folder or organize them into albums. They could save searches (in their account) by selecting the save search option. Users could revise searches or narrow searches down by adding additional terms to the original query. And they could easily repeat a new search by clicking on the "Start new search button," which deleted all the keywords users applied to their previous search and start anew.

The only search query builder the content digital museums had was autocomplete/autosuggest. Art UK and WikiArt had robust system that produced a dropdown menu (when the users started to type), which allowed users to choose options from multiple fields (artist, title, additional title information, venue, collection, people, places, events, tags, and/or acquisition method). Discover Islamic Art had an autocomplete/autosuggest query builder attached to the search box (the dropdown menu contained the words found in the glossary). They did not have spell checkers, phonetic tools, or natural language tools. Advanced query language was only available in the main search system in Discover Islamic Art and in the Art UK search system. These both supported Boolean operators, but did not strip out stop words. For example, in Discover Islamic Art, “mosaics” returned more results than “the mosaics” (39 vs. 11).

The content that was indexed for searching was the object metadata. Items that were indexed included destination pages, indexing by topic, and indexing by recent content. Art UK had a rich vertical search system, the different section (Artworks, Artists, Venus, Events, and Shop) provided users with different filters depending on the page they were searching (see above in search interface). It was important to note that while the filters were specific to what the users were searching for (e.g. for artists the user could select their nationality) if the user did not use the filters then the search system searched all the content.

Vocabulary Systems

All three of the content museums contained metadata – descriptive, structural, administrative, and embedded within the HTML. These website also contained different objects, each with their own metadata fields (events, venues, artwork, artists, etc.), though those were mainly in the Art UK digital museum. The metadata in these sites were indexed for searching (both within the site and for outside search systems, e.g., Google), to enrich the user experience. The most common descriptive metadata for artwork objects (across all three of the content digital museum sites) was title, description, style, and keywords/tags. Discover Islamic Art had the least amount of descriptive metadata (only title and description). In Art UK and WikiArt the artist descriptive metadata included nationality, art movement, and description. WikiArt had further descriptive metadata including influenced by, influenced on, friends and co-workers, family and relatives, Teachers, Painting school, and Genre. The other items in Art UK all had similar descriptive elements as the artwork objects (title, description and keywords/tags). Though the events in Art UK include title, location, suitable for, and audience.

The structural metadata for the artwork was similar across all the content digital museums (though the terminology differed). The structural metadata fields were medium (also called materials/techniques) and measurements (also called dimensions). The Discover Islamic Art digital museum was the site that had the different terminology. In Art UK the shop items had additional structural metadata including fields, print size, paper finish, frame finish, mounts, mount style, and top mount (for prints).

The administrative metadata had far more variation between the content digital museums (this was to be expected considering this type of metadata was effected by the business they belonged to). The most common administrative metadata for the artwork was artist/creator, date, location, period, provenance, license, and photo credit. Discover Islamic Art expanded on the provenance field and included when the date and origin were established, how the object was obtained, and how provenance was established. The administrative metadata for artists in WikiArt included born, died, active years and URL metadata values. In Art UK the field included associated artwork, associated venues, and associated stories. Within the Art UK shop the administrative metadata also included price.

The embedded metadata varied between the content digital museums. In Discover Islamic art, the only embedded metadata was description and keywords (but only for the overall site, not specific objects). Open graph metadata was found in Art UK and WikiArt, including: og:site_name; og:type; og:url; og:title, og:description; og:image:width, og:image:height, and og:image. Art UK only had the open graph image elements. The WikiArt also had descriptions embedded in the HTML <meta> tag (for the objects). For example, *Surprised Nymph* by Edouard Manet has the description, "Surprised Nymph, 1861 by Edouard Manet. Realism. Mythological_seo Nacional de Bellas Artes (MNBA), Buenos Aries, Argentina." The "_seo" at the end of mythological stands for "Search Engine Optimization" (Fishkin & Moz, 2015).

All three of the content digital museums had a controlled vocabulary. Art UK and WikiArt had examples of authority files (e.g. artist names) and classification schemes (e.g., artwork genres). Discover Islamic Art had classification schemes (e.g., Period/Dynasty) and was the only digital museum (in the entire study) that had a detectable synonym ring. For example, a search for "Vase" returned "Pots" and "Vessels" (and the metadata of those items did not contain mentions of the word "Vase"). Additionally, a search for "Bowl" returned "Dish" and "Plate."

Both Art UK and WikiArt had authority files for artist names. In Art UK the authority files for the preferred artist name could be found in the autosuggest features (teaching the user the correct artist name). Looking at VIAF, the artist authority files did not follow one of the standards, though there were some similarities (Pablo Picasso was similar to LCC), but many did not follow that standard (Michelangelo (1475-1564) was "Michelangelo Bunoarroti, 1475-1564" in the LCC). This was also true in WikiArt, the reason it was determined that these authority files did not follow LCSH (or any other standard for that matter) was due to the punctuation in the authority file fields. For example, the WikiArt artist name was "Fra Angelico (c.1395 - 1455)," but in LCSH it was "Angelico, fra, approximately 1400-1455" or "Raphael (1483-1520)" (WikiArt) vs. "Raphaël, 1483-1520" (LCSH).

All the evaluated content digital museums had classification schemes. These examples were found in the filters and/or facets within those sites (either searching or browsing facets). In Art UK there was a classification scheme for artwork styles (art movements), which included Art Nouveau, Cubism, High Renaissance, French Realism, Kitchen Sink, Mannerism, etc. There was also classification schemes for type (artwork) (portrait, landscape, abstract, etc.), topic (arts and entertainment, people, sport and leisure, etc.), nationality (Greek, French, Hungarian, Iraqi, etc.), type (venue) (castle or defenses, library or archive, public building, etc.), type (event) (play, concert, festival or fair, etc.), and audience (any age, especially for children, family friendly, and not suitable for children). In WikiArt the classification schemes included nation (Dutch, French, Guatemalan, Thai, Emirate, etc.), copyright country (Cuba, Japan, Canada, Belarus, etc.), art movement (Byzantine Art, Funk Art, High Renaissance, etc.), field of art (architecture, enamel, calligraphy, graphics, etc.), genre (artist) (illustration, mosaic, urushi-e, allegorical painting, etc.), school or group (Bengal school, Big Five, Group of Seven, etc.), style (Cubism, Celtic, Native Art, Hyperrealism, etc.), media (canvas, frottage, hologram, ceramics, etc.), and genre (caricature, self-portrait, tapestry, trompe-l'oiel, etc.). These were metadata fields that used predefined terms that the editor could choose from when filling out the metadata.

All three of the websites used thesaurus structures to support browsing and searching. However, none of the content museums used pre-existing controlled vocabularies (that could be determine), they were compared to the Library of Congress Subject Headings, Art and Architecture (Getty), the Humanities and Social Science Electronic Thesaurus (HASSET), etc. The WikiArt digital museum had terms added by users, which could be selected from a predetermined list or they could suggest new terms. Additionally, as users looked through the list of tags there were some inconsistences that would not have happen if they followed a standard (e.g. Shakespeare – "Mid Summers Night Dream" vs. Shakespeare – 'The Tempest' – note the different use of quotes). Users could get a sense of the current thesaurus structure by looking at all the second-level categories. For example,

Post Renaissance Art
- Baroque
- Rococo
- Neoclassicism
- Academic Art
- Romanticism
- Realism

Subgenres could then be organized beneath them (note this was an interpretation of the index terms). For example:

Post Renaissance Art
- Baroque
  - Painting
    - Religious painting
    - Allegorical painting
    - Mythological painting See High Renaissance
  - Sculpture

The list of topics let users see how Art UK organized their keywords (and their thesaurus of terms could be interpreted from this). The topics were keywords organized beneath headings (e.g., Animals and plants, Religion and belief, Ideas and emotions, Literature and fantasy, etc.) and those could be organized into a hierarchy with semantic relationships (see example below).

Animals and plants

- Animals, domestic
  - Cats
  - Dogs
  - Fish see Animals, wild
  - Birds see Animals, wild
- Animals, farm
  - Horses
- Animals, wild
  - Fish
  - Birds
  - Insects
- Plants and flowers
  - Trees and shrubs
  - Fruit and vegetables

The Art UK also used a folksonomy organization; the site allowed users to tag the artwork with their own keywords. This resulted in some inconsistencies, for example, “Nymph,” “Nymphs,” and “Water Nymphs” all described a similar keyword, but these did not overlap. Discover Islamic Art glossary terms could be interpreted from the terminology used in the site. The glossary not only had a definition of the term, but there was also a “See also” section that established relationships between the terms much like a thesaurus. Below is an example of the possible thesaurus structure of Discover Islamic Art.

Decorative Arts

- Calligraphy
  - Ahar (paper-treating process, used for Calligraphy)
  - Farsi (2. Cursive Calligraphic style)
- Ceramics

Patterns

Arabesque (Two-dimension ornamental pattern)

Ataurique (Stylized plant motif)

Azulejo (Glazed tiles in a larger pattern to cover walls)

Graffiato (scratching design)

Types

Ataifor (Bowel or deep plate)

Blue-and-white (Ceramics that appeared in Mesopotamia)

Fritware (A designation for a ceramic body consisting mainly of quartz)

Methods

Glaze (applying a thin layer of glass to ceramic before firing)

Green and Manganese (decorative ceramic glazing)

This was an attempt to show the reader the terms and possible thesauri used in the sites. All of the hierarchy structures were interpreted.

There were semantic relationships between the terms – hierarchy and associative. Equivalence relationships were harder to determine in the websites (often seen by "use for"), this relationship could only be determined in Discover Islamic Art (through the glossary). The thesaurus above showed the hierarchy relationship between the parent-child relationships (though these were interpreted). The equivalence relationships were used throughout the glossary. For example, the term "Quran" (the sacred text of Islamic Revelation) had the synonym or alternate spellings within its definition examples. The Associative relationships were stated at the end of the definition (e.g., for "arasta" there was "see also" bazaar, bedesten, etc.). The semantic relationships in Art UK were established throughout the controlled vocabulary. For example, classification schemes and authority files for regions, styles, type and topic were all subsections/metadata for the artwork. They had associative relationships because they were associated with the other categories (through the overall connection to the artwork). Additionally, as shown above in the (created) thesaurus, the hierarchy relationships could be established (e.g., Horses were a child category beneath Animals, farm for example). For WikiArt there were semantic relationships within the controlled vocabulary. There were hierarchy and associated relationships between the thesaurus terms (at least as the index terms were constructed

– shown above). This was the hardest part of evaluating the controlled vocabulary of these sites, because it required so much interpretation.

Only Art UK had a fully functioning faceted classification feature (being able to select more than once facet in a search and/or browse system). In Discover Islamic Art users could browse the content under "Permanent Collection." They could choose to browse the collection by Country (Greece, Turkey, Algeria, Egypt, Jordan, Kuwait, Palestine Territories, etc.), Period/Dynasty (Abbasids, Almohads, Atabergs, Crusaders in the Islamic world, Fatimids, etc.), start date (400 AD – 1900 AD) and end date (see image). It was important to note that once users selected a facet it took them to the collection immediately (they could not select more than one). Upon entering the collection users could select a new facet (but it would override what they had already selected – meaning that users could only ever browse using one facet). Art UK (as mentioned above) had filters/facets found in each of the search interfaces (which depended on the page the users were in). For example, these filters were found in the artwork search interface: region (Channels, England, Isle of Man, etc.), type (portrait, landscape, abstract, etc.), topic (arts and entertainment, people, sport and leisure, etc. – see more in thesaurus below), style (Art Nouveau, Cubism, High Renaissance, French Realism, Kitchen Sink, Mannerism, etc.), and license (creative commons, etc.). WikiArt did not have an overt faceted browsing system, but users could browse the artists and artworks by single facets when they selected to view objects beneath a school and group, style, etc.

Learning Digital Museums

Organization Systems

The three learning digital museum also all used a hybrid organization scheme. The top-level categories organize information beneath topical (all three learning museums), audience (Smithsonian Learning Lab and SHOW.ME), task (Web Gallery of Art), and format (Web Gallery of Art) organization schemes. For example, the top-level categories included Postcard (topical – Web Gallery of Art), Create (task – Smithsonian Learning Lab), and Teachers (audience – SHOW.ME). Exact organization schemes could be found in all levels of the digital museums. For example, the blog posts (under "Updates") were organized by date (newest to oldest) – chronologically. In SHOW.ME users could choose to access content beneath the focus entry point filters by either type (format organization scheme) or tags (topical organization scheme). Finally, in the Web Gallery of Art the Glossary items were all organized alphabetically, except for "popes in the 12th-18th centuries," which was organized chronologically.

The organization structures found in the learning digital museums were all hybrid structures. This included polyhierarchical top-level (in the Smithsonian Learning Lab and Web Gallery of Art), database-oriented structures for the content, focused entry points to access content, and subsites (Smithsonian Learning Lab and Web Gallery of Art). The polyhierarchical structure in the Smithsonian Learning Lab applied to the "My Learning Lab" (the user account) and "Smithsonian Learning Lab" (let you access the company information). For example, beneath the "Smithsonian Learning Lab," FQA was both a third and second level category. All three of the learning digital museums had focused entry points and database-oriented organization structures. For example, in the SHOW.ME site, the focus entry points (an access point that takes users directly to the content of the site) were available beneath the category "Everything" (and included format and topical categories) (see Figure 4). The database-oriented structure provided access to and organization of the content in the site. For example, in the Web Gallery of Art, users accessed content by browsing (located beneath the "Artist" category) and searching. The Web Gallery of Art database oriented structure presented the content in a table grid with an image, artist, title, date, medium, size, location, contextual link (to other works by the same artist), file size, and a link to more information (unless that was not available). The Smithsonian Learning Lab (Terms of Use, in the utility navigation) and the Web Gallery of Art (in the Tours category) both used subsites.

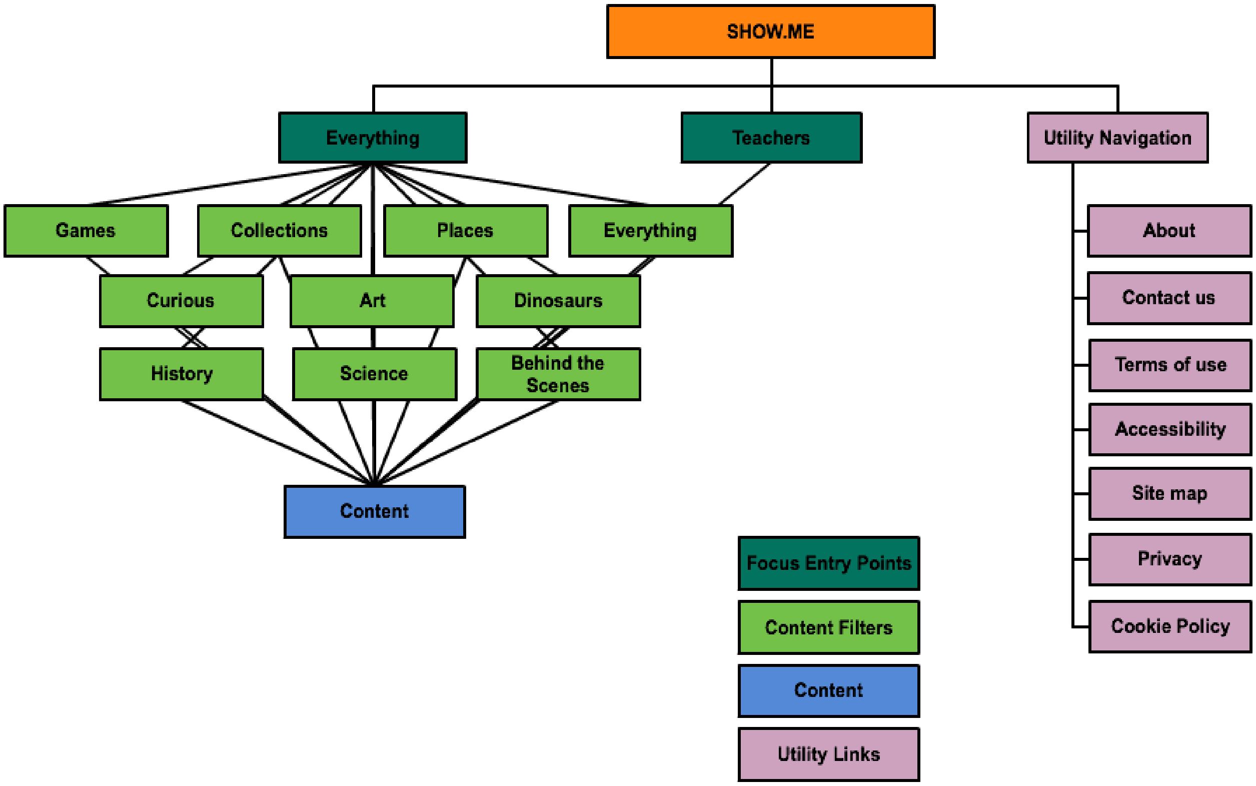


Figure 5. Diagram of SHOW.ME digital museum. This figure displayed the organization system and labels used within the top three levels of the SHOW.ME digital museum.

Labeling Systems

The learning digital museums used all the types of labels (contextual link labels, headings, navigation labels, index terms and icon labels). In SHOW.ME, the contextual link labels included extensive inline links; the labels were represented by chunks of the text. The topical/format section determined the colour of the label. For example, inline links organized under the “Dinosaur” facet meant the text was green. In Web Gallery of Art the majority of the contextual link labels were sections of the text (in line hypertext links). These were generally artist names, historical figures, art terms, and art titles. The Smithsonian Learning Lab used contextual link labels to let users know about related content. This occurred in specified areas (e.g. the right side of the page within the blog) or within the text (inline hypertext links). There was actually not that many contextual links in Smithsonian Learning Lab, items were joined together in collections rather than by links.

The heading link labels were found throughout the three learning digital museums. The Smithsonian Learning Lab used headings in the help section of the site. The two categories “Getting Started” and “For Teachers” did not link to other areas on the website, instead they organized the sections found below them. In Web Gallery of Art heading labels were used extensively in the site. For example, at the top of the content grid lists there were heading labels, which explained the content that was found in that site (Preview, Picture Data, File Info, and Comment). The “Comment” option was a labeling issue, because comments were not found there. Instead it contained links to the extended metadata field (if the object had one). In SHOW.ME there were headings at the bottom of the some of the content. In collections there were heading labels that differentiate the different sections of text. For example, the heading labels were used to separate the “Intro,” “Museum’s Description,” and “Teacher’s Notes.”

The navigation labels appeared in all three of the learning digital museums, from the global navigation down to contextual links. To view the Smithsonian global navigation labels users needed to click on the sun icon (found in the top right corner). This opened up the menu, which contained two sections – “Smithsonian Learning Lab” and “My Learning Lab.” The titles of the content in the Smithsonian Learning Lab were also navigation labels. This was how users opened the page to view and add content to collections (etc.). The SHOW.ME topics/format filter options were all navigation labels (beneath the focused entry point, “Everything”). When users clicked on them, they were taken to a page that contained content that had those tags. The Web Gallery of Art had navigation labels beneath the Artist category and at the bottom of the page (throughout the site, including under Dual Mode) represented by the letters A-Z (for example). These connected users to a list of artist names (related to whatever letter the user selected); those names were additional navigation labels that connected users with the artists’ artworks.

The index term labels found in the three sites were represented by the metadata (keywords, tags, etc.). The SHOW.ME digital museum used index term to filter the results. These were tags (general and teaching specific) and well as topics. The tags were both browsable (by popularity within a topic) and searchable. Web Gallery of Art had many different index terms labels throughout the site (embedded keywords, glossary terms, facets for browsing, etc.).

The glossary for art terms (accessed below “Glossary” and “Database”) included index term labels, which were then defined (e.g., “original sin” – evil was transmitted to mankind after Eve ate from the Tree of Knowledge). The Smithsonian Learning Lab index terms were labeled as keywords, tags, subjects, etc. For example, the keywords associated with the “Arrangement in Black: Portrait of F. R. Leyland” by James Whistler were “portrait,” “United States” and “American Art” (these index terms were also contextual link labels, which took users to a new search results page where the content contained that keyword).

There were icon labels in all three of the digital museums, especially in the Smithsonian Learning Lab, which used icons as the main labels of the sites. The Smithsonian Learning Lab used an icon label for the global navigation menu, located in the top right corner (found across the site design). A sun represented the global navigation menu (the Smithsonian’s logo). The colours of that icon label changed depending on what page the user was in. The Smithsonian Learning Lab described the icon labels used in the help page. The help pages divided the icons into “types” (resources, content type, content actions, etc.). SHOW. ME used icon labels to differentiate the types of content (Collection, Website, Games, Stuff to Read, Video, Events/Exhibitions, and Place). This site also used icons that let users share the content (Facebook, Twitter, Pintrest, email, and print). The Web Gallery of Art used the least amount of icon labels. There was only one use of an icon label (and it includes a lower case “i”) in the site. These were found in the content and under the “Comment” heading. When users clicked on this icon it took them to further information about the content it was associated with.

The three learning digital museums also had labeling consistency issues. The Smithsonian Learning Lab had issues by relying almost entirely on icon labels (which were hard for users to learn). Though the help section described the icons to offset this issue, during the first evaluation the website the icons were confusing (e.g., the flag, did it mean flag something the users wanted to return to or a problem in the site – it meant the latter). These should only be used if the site had users who were willing to spend time learning them (which they might, but the site should remember that not all teachers – their target audience – are tech savvy). The second issue the Smithsonian Learning Lab had with labeling consistency was the label given to their blog, “Updates.” The first few times this digital museum was evaluated, it was assumed that this section referred to technology updates for the site not the blog (this issue was subjective). SHOW.ME labeling consistency issues were mainly due to the capitalization of the topics/tags. The topics organized beneath "Everything" were all capitalized and the topics stated within the metadata field only had the first letter capitalized (a presentation issue). There were also capitalization issues in the tags. For example, the tag "crime and punishment" (all letters lowercase) was mixed with "Nautical Heritage," "History," and "Pirates." There were also issues with mixing different types of labels (under "Everything" there were types mixed in with topics even though the topics could be sorted by type). The topic labels used to browse the site had issues with granularity. Some of them were broad (Art, History, Science, etc.) but they were mixed in with more specific terms (Dinosaurs). The Web Gallery of Art labeling inconsistencies included those between the navigation labels and the headings on the pages that they went to. For example, the tour navigation label called "French art in the 15th – 18th centuries" went to the page where the heading was called "French Painting and Sculpture in the 15th-18th centuries." Also in this site, labels that linked to the homepage were called different things in different places. This inconsistency in label terminology was also reflected by the "Time-Frame" (simple search interface at the bottom of the page) vs. "Time-line" (advanced search system) labels, both filters limit the search by fifty-year date ranges. There were also issues with capitalization. The global navigation labels were fully capitalized, most of the other labels had the first letter capitalized, but all the labels beneath the "Info" tag were lowercase.

Navigation Systems

These three digital museums had navigation systems, but not all of the main types were represented (none of them had breadcrumb navigation). The global navigation in SHOW.ME was the focused entry point filters. Arranged at the top of the page (initially organized beneath the "Everything" heading), the drop-down menu and the filters were available on every page in the site. The top-level heading (originally "Everything") changed depending on the filter the user was in (e.g., Dinosaurs, History, Teachers, Search, etc.). Web Gallery of Art global navigation was found throughout most the site, though it was not present in the subsites and the "Dual Mode" (which prioritized screen space for the comparison of pages). The global navigation categories were "Home," "Mobile," "Artists," "Search," "Tours," "Dual Mode," "Glossary," "Music," "Postcard," "Database," "Sources," "Guestbook," "What's New," "Email," and "Info" (emphasizing breadth over depth). In the Smithsonian Learning Lab the global navigation was found in the upper right corner of the site, represented by a sun icon. This was where registered users accessed their account and found the help documentation. This global navigation emphasized depth over breadth in the polyhierarchical structure of the site.

Local navigation in the learning digital museums varied from site to site. The SHOW.ME digital museum did not really have a local navigation system; the closest thing to it was the list of results found when users selected one of the focused entry point filters (including search). They included images, navigation labels, and content metadata. The Smithsonian Learning Lab local navigation was found beneath the global navigation categories, especially in the "About" and "Help" sections of the website. For example, under the "About" page there was ten additional pages (local navigation) that covered information like blog access (under "Updates") or how to volunteer at the Smithsonian ("Volunteers"). Also in the Smithsonian Learning Lab there were three focus entry points presented on the homepage, which let users access the collection. These were a type of local navigation because they were only accessed on the homepage, but they represented the main actions users (teachers or students) could do within the site. The Web Gallery of Art had many types of local navigation. For example, in the "Info" category there were nine different second level categories accessed using the icon label. There was local navigation located on the left side of some of the pages. In the Tour homepage there was a list of tours available, with the associative numbers and shortened titles.

All the learning digital museums had contextual/hypertext navigation links. The Web Gallery of Art used a lot of contextual/hypertext navigation features, especially inline links. There were many different inline links throughout the site, represented by blue text labels found in the content of the site. These commonly represented links to artists, other works, and art terms (found in the glossary). Additionally, when users conducted a keyword search the list of results had contextual links to other works of art by that artist. Those contextual link labels were called "Other works by the artist…". In the SHOW.ME digital museum there were inline links and links to similar content organized at the bottom of the page under the heading label "You Might Also Like." They linked together items that had similar topics, similar content, or similar types.

Finally the Smithsonian Learning Lab contextual/hyperlink navigation was used throughout the site (though less than in previous museums). For example, when users viewed an item they could see the collections that contain that resource. There were also inline contextual links, which were most often found in the "About" and "Help" pages. They used the colour design, highlighting the linked text blue.

Utility navigation in all three of the learning digital museums was found in the expanded footers of each site. The SHOW.ME digital museum used the utility navigation for the legal and technical information related to the site. The utility options included "About us," "Contact us," "Terms of use," "Accessibility," "Site map," and "Cookie Policy. " The Smithsonian Learning Lab utility navigation options included links to the "About," "Contact Us," and "Privacy" pages. The "Kids Online Privacy Statement" linked to the Smithsonian site. There were also links to social media (Facebook, Twitter, and Google+) represented by icons. The Web Gallery of Art utility navigation options included easy access to the A-Z artist index, as well as "quick links" to pre-organized content – Medieval, Decorative, and Architecture items. The Web Gallery of Art utility navigation also featured a search interface.

The supplemental navigation features found in the three learning digital museums included a sitemap, indexes, guides, control panels, toolbars and pagination navigation. The only supplemental navigation feature in the SHOW.ME learning digital museum was a sitemap. The information in this sitemap did link to filtered content, but it did not represent what was available on the main page. For example, the options listed in the site map – "Tudor," "Places," "Animals," and "Ancient Egypt" could only be accessed using search or clicking on those tags. The Smithsonian Learning Lab had a few more supplemental navigation features. There were types of guides found on the website, collections were made up of several individual content items (most often organized by users). Users needed to click through the collections (like lesson plans or power-points that users could present to students) in order to teach the subject. For example, the collection "Early American Nationalism" has 15 items within it. On the front page of the collection, all 15 items were shown as thumbnails, but once users entered the collection they had three navigation options (the previous page, next page, and back to the collection homepage – by clicking on the "four square" icon in the lower right page). The Smithsonian

Learning Lab also had control panels (when editing profile information or adding resources), toolbars (when looking at the content in the site – options included viewing the metadata, adding the content to a collection, sharing the content with social media, downloading the content, etc.) and pagination navigation (at the bottom of the search results). The Web Gallery of Art also had many supplemental navigation features. These included A-Z indexes (for artists, art terms, famous families of the Italian Renaissance, etc.), guides (in the form of tours, organized around time periods and locations), control panels (found in the postcard section – when users wanted to send a postcard they had to fill out information like the recipients email address, etc.), toolbars (these were found in the larger image popups and gave users the option to change the background colour, change the size of the images by percentage, and fit the height or width of the image to the popup size), and pagination navigation (available when the user browsed/searched for content).

The learning digital museums each had some sort of advanced navigation features. The Web Gallery of Art only had one type of advanced navigation feature – visualization. In the tour "Overview of Italian Painters from 1200-1750" users could navigate using a map of Italy (users clicked on the different locations to see the associated art). The SHOW.ME advanced navigation features included visualization and social navigation. Every content item (accessed via browsing or searching) had images associated with them. Users could use the pictures to help them decide what to select. This site also used social navigation. The tags (applied by the website) let users navigation between content. For example, if users selected the tag "19th Century (1801-1900)" they were taken to a new page that had eight different content items (all containing that tag). The Smithsonian Learning Lab had three advanced navigation features – customization, visualization and social navigation. The site emphasized the customization of information to suit the user needs because users were the driving force behind the website. They could upload, connect, and create their own content. The Smithsonian Learning Lab let users copy collections and then edit or add information, customizing them for their needs and (if teachers) their students. The visualization feature referred to how the search results presented the images with metadata (as well as the extensive use of icon labels). The Smithsonian Learning Lab also used some aspect of social navigation. Users added much of the metadata for the collections (and the objects they uploaded). This metadata was used to navigate to the collection, though users did not seem to be able to add tags. This may be because they needed to be properly formatted (the site used the Library of Congress Subject Headings and Authority Files).

Search Systems

The Smithsonian Learning Lab and Web Gallery of art had one search system each (though each search system was accessed in multiple locations). The SHOW.ME digital museum had two search systems, one was an example of vertical search, it only searched events, exhibitions, and venues (by location or postal code). Both search systems in SHOW.ME and the Smithsonian Learning Lab prioritized recall over precision. For example, a search for "Dinosaur Bones" in the Smithsonian Learning Lab returned the results "Australia Used to be a Haven for Giant Penguins." The Smithsonian could have emphasized recall over precision because of the facets available for narrowing results. The Web Gallery of Art search system prioritized precision over recall. The search interface was set up so there were exact matches (though the use of stemming extended the results when searching text/title).

The search interfaces for the learning digital museums varied in complexity. The Smithsonian Learning Lab had one search system and four different interfaces. The simplest versions were found under a persistent link located in the top left corner of the page and on the homepage. Both had a simple search box and magnify glass icon button (as did the interface under "Share"). The three other search interfaces were found under the Discover, Create, and Share focused entry points (accessed on the homepage). Under "Discover," the search box was shortened (still with the magnify glass icon) with filters to pre-set search results ("Trending" – sorted by popularity and "Surprise Me!" – randomness). Under the "Create" focus entry point, the search interface was longer (still with the magnify glass icon button) with the text "Find a Collection to Adapt and Customize." There was also a box (on the left) to let users start a new collection (link to "My Learning Lab > Start New Collection" page). Under the "Share" focused entry point was a simple search box and magnify glass icon button. Those three search interfaces were all located below the top banner in the center of the page. The Web Gallery of Art had two search interfaces (simple and advanced, see below). The simple search interface had two search boxes. One was for an artist's name and the other was used to search the descriptions of the items. There were two dropdown menus – Time-frame (periods of 50 year increments from 0701-1900) and Form (painting, sculpture, architecture, ceramic, etc.). Finally, there were two buttons "Search" and "Clear" (to clear the form). The two search interfaces in SHOW.ME were differently designed. The site wide search interface was located in the top right corner of the website (throughout the site). When users clicked on the magnify icon, the search box expanded across the page. This was a simple search box with a magnify glass button. The second search interface (for the events, exhibits and venues) was very simple. Located below the banner, but above the list of events/exhibits/venues, it was a simple search box and a magnify glass icon button.

Only Web Gallery of Art had an advanced search feature, which was very complex. The first section had a search box for artist searches with a dropdown option that contained every artist name found in the A-Z index. There were two more search boxes (to search in the title and description metadata). There were then four dropdown fields – Time-line (periods of 50 year increments from 0701-1900), School (American, Catalan, Flemish, Greek, etc.), Form (painting, mosaic, tapestry, etc.), and Type (religious, mythology, landscape, etc.). Then there was another search box for the location field and a dropdown option (A-Z list of possible locations). Finally, there were two search buttons – "Search" and "Clear."

How the search results were displayed depended on the search system. The Smithsonian Learning Lab displayed content in multiple ways. Firstly, results could be classified either as "Resources" or "Learning Lab Collections" – selected by clicking on tabs at the top of the page. Users could view results either in a list or grid pattern (both only presented 24 per page). The information displayed in the list format included an image, title, location, and icons (share, favorite, add to collection, and file type). Learning Lab collections also contained the collection creator information and the number of items in a collection (an icon) metadata. The grid (thumbnail) view showed information when the user hovered their cursor over the image – title, location, and icons (share, favorite, add to collection, and file type). The Web Gallery of Art presented the search results in a table/grid pattern (20 per page). The information displayed included an image, picture data (artist, title, date, medium, size, collection and links to other works by the artist), file information (image size, colour, and file size), and "Comment" (aka link to further information). The SHOW.ME search results were displayed in the same way for both search systems, in a grid/thumbnail pattern (with 16 per page). The information displayed depends on the type of object. Most contained an image, title, type, and description. Events and exhibits also had the date, location and distance from the town/postal code searched.

All the search results were either ranked and/or sorted in the learning digital museums. The Web Gallery of Art only used one sorting option, alphabetically by artist (no matter what was searched). The SHOW.ME digital museum search results were either ranked by relevance (the main search system) or by distance (for the events/exhibitions/place search system). The distance ranking sorted the closest first, getting farther away as the user scrolled down the list. The Smithsonian Learning Lab search results were initially ranked by relevance (best matches) or popularity if selected, but they could also be sorted alphabetically (Title A-Z and Title Z-A). Learning Lab Collections could also be sorted chronologically (Date Added (Oldest First), Date Added (Most Recent First), Date Modified (Oldest First), Date Modified (Most Recent First)).

The Web Gallery of Art and SHOW.ME (both systems) only had one additional action. The search interface was easily accessed (to start a new query) in both those digital museums. In the main SHOW.ME search system it seemed like users could use topics to narrow the results down, but selecting one of those types/topics instead took the user to a brand new page (the same page they would get if they had selected it from the homepage). The Smithsonian Learning Lab had the most additional actions available. These included saving a subset of results (by adding them to a list of favorites) and narrowing results down using filters. In resources, the filters were Resource Type, On Exhibit, and Resource Provider. Learning Lab Collection filters included Collection Type, Item Type, and Annotation Type. Users could also repeat/new search by using the search interface (at the top of the results), which was populated with the search term.

Some of the learning digital museums had query builders and advanced query languages available. The SHOW.ME digital museum (both search systems) did not support phonetic tools, stemming, spellcheckers, etc., though it did offer users a drop down list of previous search terms. Those search systems also did not support Boolean searching or strip out stop words. The Web Gallery of Art search did not have a spell checker, phonetic tools, or natural language processing tools. It did have some stemming (a search for mythology returns results that just had “myth”) and minimal autocomplete/autosuggest (past search terms could be selected from the dropdown field). This search system did support query languages, it stripped out stop words, and Boolean language was the default when the users entered two terms (in the title and text fields), the system read it as “Term” AND “Term.” The Smithsonian Learning Lab query builders included stemming (a search for paintings returned results with “paint”) and natural language processing tools (results were returned for the search “What are the parts of an airplane,” though it returned way more results without a question mark – 98 vs. 6). The Smithsonian Learning Lab did support Boolean languages and strips out stop words (the, a, and, etc.).

The content indexed for searching depended on the search system, though they all indexed the content metadata (this included indexing by topic) and destination pages. This was the only way that the Web Gallery of Art was indexed. The Smithsonian Learning Lab also indexed recent content (which was why users could sort using this option in Learning Lab Collections). The search zones for the Smithsonian Learning Lab were resources and collection (search results were sorted into those two categories). The main search system in SHOW.ME also included full-text indexing (in addition to indexing metadata and destination pages). In the SHOW.ME events/exhibitions/place search system the location of the content was indexed (an example of vertical searching).

Vocabulary Systems

All three of the sites had metadata – descriptive, structural, administrative, and embedded (in the HTML <meta> tag). Different content had different metadata fields. The descriptive metadata that most of the content had was title, description, and keywords/tags. The Web Gallery of Art descriptive metadata (for images) only had title and description. The descriptive metadata for artists included period, school, and description/bibliography. The descriptive metadata for Smithsonian Learning Labs also included note (title and category), which was true for all the items in that digital museum (though what the note contained differed – e.g., lesson plan notes included education use, learning resource type, feature, and interactivity type). The notes section for the content in the Smithsonian Learning Lab also contained structural and administrative metadata as well (see below). The Smithsonian Learning Lab collections contained additional descriptive metadata – subjects, age range, and education features. The descriptive metadata in the SHOW.ME digital museum depended on the content (collections, stuff to read, events/exhibitions, games, videos, websites, and places) though they all had a title, description and tags. Additional descriptive metadata includes teaching tags (in stuff to read, collections, etc.), and topics (videos and websites).

The structural metadata described the make up of the objects (both in real life and digitally). The structural metadata for the learning digital museums included medium/physical description, and type. In Web Gallery of Art the structural metadata included medium, height/size, image size, file colour, and file size. The Smithsonian Learning Lab structural metadata most often included type (e.g., collection, website, video, etc.), physical description, and notes (dimensions, duration, and/or time required). The structural metadata in SHOW.ME just had the type.

The administrative metadata often contained the date and creator/artist information, but across the learning digital museums this metadata field differed. The SHOW.ME administrative metadata for collections included artist, date, held at, and production place. The administrative metadata for events/exhibitions, games, videos, and websites included from (the place this object was located at) and link (the URL for that object). The Smithsonian Learning Lab administrative metadata was (for images) creator, date, identifier, view original, additional info, and notes (contained within and contact information). Other Smithsonian Learning Lab administrative metadata included view original (audio, video, text documents, lesson plans, and resources), additional info (videos, text documents, lesson plans, and resources), and notes (copyright, views, citation, and accessibility control). Learning Lab Collections also had the administrative metadata – adapted from, last modified, created by, date published, and published by. The Web Gallery of Art administrative metadata (for images) contained creator/artist, date, location, artist birth/death (date and location), and catalogue number. For artists, the administrative metadata included artist name, and born/death date.

The embedded metadata ranged from simple (just a description and keywords) to complex (open graph and twitter metadata). The Web Gallery of Art and SHOW.ME had had title, description, and keyword embedded metadata. For example, the painting (in Web Gallery of Art) *Perseus and Andromeda* by Giorgio Vasari had the embedded description "Page of Perseus and Andromeda by VASARI, Giorgio in the Web _ of European painting, sculpture and architecture (700 – 1900)" and keywords "VASARI, Giorgio, Perseus and Andromeda, image collection, virtual museum, database, postcard." While these were not very specific, they were still attached to the object HTML. The embedded metadata in the Smithsonian Learning Lab included the title and description for the content. This digital museums contained open graph (og:type, og:title, og:image, og:url, og:description, og:site_name) and twitter (twitter:card, twitter:title, twitter:image, twitter:description, twitter:site) metadata. The Smithsonian Learning Lab also used the meta: item prop tags to add the name, URL, learningResourceType, EducationUse, dateCreated, dateModified, Artist, Keywords, Additional Info, and Notes metadata.

All of the learning digital museums had a controlled vocabulary, though none of them had a synonym ring (that could be determined). The Smithsonian Learning Lab artist/creator names (of resources – not collections) followed Library of Congress Authority Files. Keywords followed the Library of Congress Subject Headings (LCSH), but tags (in collections) did not follow the LCSH subjects. SHOW.ME and the Web Gallery of Art did not follow pre-existing controlled vocabularies. SHOW.ME had classification schemes, which were categorized by topic (Curious, Art, Dinosaurs, History, Science, and Behind the scenes), and type (collections, stuff to read, events, exhibitions, games, videos, websites, and places). These were the two ways the SHOW.ME website organized information (and they were cross-listed, e.g., there were collections that belonged under history, etc.). SHOW.ME did not have an authority file. The Web Gallery of Art had both authority files and classification schemes. The A-Z artist index contained the authority file for the site; this listed the preferred names of the artist. For example, a search for "Titian" did not return any results (despite the fact that this digital museum covered the time period and movement that the artist belonged to). This was because users needed to search with "Titian's" full correct name, "TIZIANO, Vecellio" in order to return the 279 artworks by Titian available on the Web Gallery of Art. This site also had numerous classification schemes (found in the search and browse filters). These predefined terms were part of the metadata (necessary in order to filter results), which included school (American, Flemish, Hungarian, Norwegian, etc.), form (painting, sculpture, graphics, illumination, architecture, etc.), type (religious, historical, mythological, landscape, portrait, still-life, etc.), period (Medieval, Early Renaissance, Northern Renaissance, High Renaissance, Mannerism, etc.), and profession (painter, sculptor, graphic artist, miniaturist, illuminator, etc.).

The Smithsonian Learning Lab used the Library of Congress Thesaurus for Graphic Materials. Though there were some inconsistencies in the construction of the thesaurus terms. For example, the keyword “Performing arts – Dance – Ballet” also appears as “Performing arts/Dance/Ballet.” SHOW.ME did have a searching/indexing thesaurus. Tags (keywords) were indexed for browsing/searching, but these did not come from an established standard (the inconsistencies in the tags – capitalization, etc. – meant that they were most likely applied by hand). The Web Gallery of Art used a thesaurus (shown by the glossary of art terms and the classification schemes). For example (from the glossary):

- Architectural Components
  - Arch
    - Arch Components
      - Arcade
      - Architrave
      - Archivolt
      - Blind
      - Corbel
      - Flying Buttress
        - Frieze
      - Keystone
      - Pinnacle
    - Types of Arches
      - Horseshoe arch
      - Tudor arch
  - Column

Column Components

- Abacas
- Base
- Entablature
- Fluted
- Trumeau
- Pediment plinth

Orders of architecture (Types of columns)

- Ionic Order
- Doric Order
- Corinthian order

This was an example of the possible thesaurus that existed in the Web Gallery of Art (created from the art terms glossary). These terms were used throughout the text when discussing architecture, artist biographies, and metadata descriptions.

The semantic relationships depended on the learning digital museum. The Smithsonian Learning Lab had examples of all the types of semantic relationships (hierarchy, equivalence, and associative) that were found in the LCSH. For example, the hierarchy relationship (parent-child) could be seen here "Art – Architecture – Details – Columns." In that example "Art" was the highest-level term, with the nested (child) beneath it ("Architecture") and "Details" was a child category of "Architecture" (and so on). Equivalence relationships could be seen in variant terms (e.g., Airplanes vs. Planes(Airplanes) were variant terms). Associative terms were shown by the related terms (RT) classification. For example, Archival materials had the related term (RT) Manuscripts. The SHOW.ME digital museum tags did show some semantic relationships. For example, the terms below showed possible hierarchy and associative relationships (created from the tags in the site):

Science

- biology
- zoology

RT Nature

The semantic relationships in the Web Gallery of Art included hierarchy and associative relationships found in the glossary (art terms, etc. – see above). There were associative relationships also found in the classification schemes of the site (organized together by the site under the search filter). For example, the terms religious, historical, mythological, landscape, portrait, still life, interior, genre, study, and other were all organized under "type." These terms were associated with each other because they reflected a "type" of artwork.

The Smithsonian Learning Lab and the Web Gallery of Art both had rich faceted classification systems. The SHOW.ME digital museum let users select either type or topic categories, but if topic was selected it could be filtered by type (be selecting the "sort" option above the results). The Web Gallery of art had facets in the browsing system (under Permanent Collection) and the search system (under Database). For example, in the Permanent Collection users could choose to filter results by school, period, time-line, and profession. They could select any and all of these options before browsing. The Smithsonian Learning Lab facets were found on the left side of the search results page. The facets depended on what category (Resources or Collections) was being searched. For example, Resources filters included – resource type, on exhibit, and resource provider. The Learning Lab Collections facets included subject, age range, education features, collection type, item type, and annotation type.

Bugs/Issues

There were twenty-two bugs found, at least one in all of the nine digital museums evaluated (see Table 2). The bugs were recorded in a Microsoft Excel spreadsheet while conducting the heuristic evaluation. The number of bugs found in each digital museum ranged from one (in WikiArt and the Web Gallery of Art) to seven (Cleveland Museum of Art). These included issues with the search systems (buttons only half appearing), navigation issues (broken links, overlapping control panels, a popup that did not appear properly, etc.), and organization issues (in an alphabetical organization scheme there was a "D" item organized above a "C" item). For example, when evaluating the databases (supporting resources) in the Web Gallery of Art, it was observed that one of the links was broken ("Directory of Online Museums").

Once the heuristic evaluation was completed, the bugs were retested to ensure that there were no false positives. During the secondary tests it was determined that there were four false positives (bringing the total number of bugs down to eighteen). The bugs found to be false positives were mostly related to the search system, for example in the Discover Islamic Art a bug was initially reported that when searching for “painting” (classified as the medium/technique) there were only 18 results, but when users clicked on page two of the search results there was suddenly 50 results. When retested, the search results and pagination navigation behaved as expected (there were only 18 results returned, and only 18 that the user could scroll through). Another false positive example was in the Cleveland Museum of Art. During the evaluation the contextual/related information metadata extended into the image frame (this happened in the page for Caravaggio’s *Crucifixion*). When tested again, this did not reoccur (the image and metadata/contextual information did not overlap). After excluding the false positives (there were four in total), each of the nine digital museums still had at least one bug associated with them. Once all the bugs were retested, they were examined to see how they could be fixed (discussed below in findings). The table below lists all the bugs found during the heuristic evaluation, including the digital museum it was found in, the bug/issue, description, heuristic it is attached to, and if they turned out to be false positives (“True or False”).

Digital Museum	**Bugs/Issues**	**Description**	**Heuristic**	**True or False**
Art Gallery of Ontario	Spoliation Research	Alphabetical organization of provenance metadata studies is out of order on the first page (A, D, B, C…). This is after the introduction link	Alphabetical organization scheme	TRUE
Art Gallery of Ontario	ShopAGO > AGO Art Rentals & Sales	Broken link (or missing page in Flickr) - When you click on "Basquait: Photographs by Roland Hagenburg" you get a "Page not found" in Flickr	Contextual Navigation	TRUE
Art Gallery of Ontario	Near "Group Visits"	There are floating categories that are supposed to belong in the #2 Visit top-level category, but they are only accessible by clicking on "Group Visit" under the top-level "Visitor Information" category in the #1 homepage.	Navigation system or Organization structure	TRUE
The National Portrait Gallery	Advanced Search	The "Portrait set" drop down menu extends off the advanced search border.	Advanced search	TRUE
The National Portrait Gallery	Search result page for "Other pages"	The list of the pages for the "Other pages" search results section runs off the page	Search results or pagination navigation	TRUE
Cleveland Museum of Art	Meet the Director and Volunteer opportunities	There is a double menu at the bottom of the page (all the same links are repeated in two sections one right on top of the other).	Contextual Links	TRUE
Cleveland Museum of Art	#paradethecircle page	Link to empty page (no content, etc.)	Navigation Links	TRUE

Cleveland Museum of Art	Follow CMA on Tumblr (under About> Connect with CMA	Link to an in maintenance/missing page in Tumblr	Navigation Links	TRUE
Cleveland Museum of Art	On the "Learn" homepage	The popup (in the lower left corner) that contains links to terms and conditions as well as a site map doesn't popup all the way	Navigation system	TRUE
Cleveland Museum of Art	On the Napoleon's Logement image page	The search "Clear" (I think) button is cut off (only half of it appears).	Search interface	TRUE
Cleveland Museum of Art	At the bottom of the page menu	If the window is too small (about 6.5 inches) the bottom menu overlaps with the hours information	Utility navigation	TRUE
Cleveland Museum of Art	In the collection object view (Caravaggio Crucifixion)	If the image is too big or when the page is expanded, the details and related information tabs extend into the image frame	Navigation links	FALSE
Art UK	On the homepage	The up icon button, located at the bottom of the page doesn't work on the homepage. When you click It, it turns red but the screen remains at the bottom of the page.	Utility navigation	TRUE
Art UK	The global navigation within the shop page	The global navigation menu changes in organization when you enter the shop site. Now the Shop information is located at the top of the page instead of the bottom.	Global navigation	TRUE
WikiArt	The control panel form used to finish "quick edits"	The control panel form used to finish "quick edits" went wonky (the text boxes doubled and overlapped, but you could still enter text in both boxes.	Control panels (supplemental navigation)	TRUE
Discover Islamic Art	The search results page (for Ghaza)	Only one result is found with the search of Ghaza, but you can't scroll to the bottom of the page (when you try it shoves the page to the top, not letting you get past a section of the page)	Search results	FALSE
Discover Islamic Art	The glossary terms are not scrollable	You can only view the full glossary list if you zoom out (and at that point you cannot read the terms anyways)	Indexes	TRUE
Discover Islamic Art	The search results	A search result for "painting" says that there are only 18 results, but when you click next or page two in the pagination navigation you all of a sudden get 50 results.	Search results	FALSE
Smithsonian Learning Lab	The Dashboard area in "My Learning Lab"	In the Dashboard page of 'My Learning Lab" the text overlaps in the list of collections under "My Collections."	Customization navigation (advanced)	TRUE
SHOW.ME	Top Ten Places to See Dinosaurs	There is a broken link in this article. The "museums website" inline contextual link leads to a 404 error	Contextual Navigation	TRUE
SHOW.ME	The Events/Exhibits page	When trying to delete content from the search bar using the backspace button the web browser goes back to past pages (e.g., to the search I made for "Sculpture")	Search interface	FALSE
Web Gallery of Art	Database	The "Directory of Online Museums" is a broken link	Navigation links	TRUE

Table 2. List of Bugs from the Heuristic Evaluation. Solutions will be suggested in the findings section below.

Findings

The selected digital museums had similar information architecture characteristics in each of the main categories (organization, labeling, navigation, search and vocabulary systems). The similarities could be seen across all types of digital museums. For example, each of the nine digital museums used hybrid organization schemes and each had examples of global, local, contextual, and utility navigation. A full summary of the similarities in the select digital museums is organized below.

All the digital museums used a hybrid organization scheme. The top-level categories used ambiguous organization schemes, generally topical and task specific. This was true for all nine of the digital museums. For example, the Art Gallery of Ontario (brochure) had the category “Exhibition & Events,” WikiArt (content) used “Artworks,” and Web Gallery of Art (learning) had “Postcard” categories (all topical organization schemes). A few of the digital museums had examples of exact organization schemes in the top-level categories. For example, the Web Gallery of art used the format scheme, “Music.” The nine digital museums commonly used exact organization schemes in the lower levels of the site (to arrange content). Content could be organized chronologically (by date of creation), alphabetically (by artist last name), and/or geographically (where the artist was from or where the artwork is/was located). For example, glossaries or A-Z indexes were organized alphabetically (these were found in the National Portrait Gallery, Discover Islamic Art, Web Gallery of Art, etc.). Art UK and SHOW.ME let users see the different venues/places that the content was in or from. The hybrid organization scheme was the criterion for digital museums – ambiguous organization schemes in the top-levels and exact organization schemes used to organize the content.

The nine digital museums all had a hybrid organization structure. The websites were designed with a polyhierarchical top-level structure (items were cross-listed so it was not a strict hierarchy) with a database-oriented bottom level used to organize content (with the metadata). Most of the sites had subsites (different sites attached to the main page using the same website design) scattered throughout the organization structure. For example, there were subsites in the

Cleveland Museum of Art (the library and archive catalogue), in the Discover Islamic Art (users account page), and the Smithsonian Learning Lab (the terms and conditions found in the utility navigation). The content and learning digital museums also used focused entry points, which provided a quick and organized way for users to access content. For example, the Art UK used focused entry points by organizing the content first under "Artwork," "Artists, " "Stories," and "Topics." Users then accessed the content by searching those sections or selecting an option highlighted on those pages. The Smithsonian Learning Lab used focused entry points organized beneath "Discover," "Create," and Share" categories (using search). The WikiArt digital museum also used a hypertext structure (often found in wikis), which used contextual links to create connection between pages (this was mainly used to connect content together by genre, media, style, etc.). All the digital museums used a hybrid structure, which provided flexibility when organization the content of the digital museums.

All the nine digital museums had examples for each type of label (contextual link labels, headings, navigation labels, index terms, and icon labels). These were used for the same purpose throughout the sites. For examples, contextual link labels were found organized beneath "Related…" headers or within the text of the site (as inline links). The inline links used chunks of text as the label and were distinguished by colour and/or underlines (which reflected the colour design of the digital museum). For example, the Cleveland Museum of Art and the Smithsonian Learning Lab both used blue inline contextual labels (because that was part of the colour scheme). The Heading labels were used as the titles and section headings within the pages. Navigation labels were used in every level of all nine digital museums (from global down to utility navigation). Index term labels were an important part of the digital museums because they connected the user with the content. Index term labels were represented by the descriptive metadata terms – keywords, tags, etc. For example, ART UK organized subject and theme keywords together, so users could browse using index terms (e.g., ideas and emotions, towns and buildings, science and knowledge, etc.). Icon labels were found in all nine digital museums, but they ranged in popularity. For example, the Smithsonian Learning Lab used icon labels as the main labels of the site, but the Web Gallery of Art only had one example. Most of the digital museums (all but Web Gallery of Art) used icon labels to represent links to social media (Facebook, Twitter, Instagram, etc.). Every digital museum had labeling consistency issues.

These also ranged in importance from minor (capitalization issues) to major. For example, the Art Gallery of Ontario labels completely changed from one page to another (e.g. "Events & Exhibitions" turned into "What's On"). A criterion necessary for all the digital museums would be a review of the labeling structure and terminology they used.

The nine digital museums all used global, local, contextual/hypertext, and utility navigation features. Five of the nine digital museums (Art Gallery of Ontario, National Portrait Gallery, Cleveland Museum of Art, WikiArt, and Discover Islamic Art) contained breadcrumb navigation. The global navigation was generally located at the top of the page (or accessed at the top, like Art UK's global navigation button located in the top right corner). The local navigation had some similarities between the nine digital museums. For example, many of the digital museums (the Smithsonian Learning Lab, Art UK, National Portrait Gallery, Cleveland Museum of Art, etc.) designed their local navigation using images and navigation labels. As mentioned above in the labeling section, all these digital museums had contextual/hypertext navigation. These were represented by links within the content (to other artworks, artist pages, collections, etc.) or inline links, which connected similar themes and subjects (e.g. by clicking on tags to see other works that contained those tags as well). Utility navigation in all the sites was found in expanded footers. These had links to the term and conditions, information about the site and social media icons.

The supplemental navigation features varied between the nine digital museums, but all of them had at least one type. Only one digital museum had all the supplemental navigation features. The Cleveland Museum of Art had a sitemap, index (for artists), guides, control panels, toolbars (for images – to share, email, or print them), and pagination navigation. The rest of the digital museums had some combination. For example, the Art UK only had control panels (used to make a donation) and toolbars (to post the image they were looking at to social media). Only two sites had a site map, the Cleveland Museum of Art and SHOW.ME. Indexes were found in the Art Gallery of Ontario, National Portrait Gallery, Cleveland Museum of Art, WikiArt, Discover Islamic Art (glossary), and the Web Gallery of Art. All the sites except WikiArt, Discover Islamic Art, and SHOW.ME had some form of guide (e.g., quizzes, tours, or for buying tickets). The digital museums that had guides were the Art Gallery of Ontario, National Portrait

Gallery, Cleveland Museum of Art, WikiArt, Smithsonian Learning Lab, and the Web Gallery of Art. SHOW.ME was the only site that did not have control panels (used to edit artwork, edit user information, etc.). Additionally, the only digital museums that did not have pagination navigation were Art UK and SHOW.ME (art was viewed in a page that loaded more results when the user reached the bottom of the page). The type of supplemental navigation found in the digital museums varied, but they all had at least one example. Here is a look at the supplemental navigation features by the numbers:

- Eight of the nine digital museums had control panels,
- Seven of the nine digital museums had pagination navigation
- Six of the nine digital museums had A-Z indexes,
- Six of the nine digital museums had guides
- Four of the nine digital museums had toolbars
- Two of the nine digital museums had sitemaps

Each of the nine digital museums used advanced navigation features. The nine digital museums have examples of only three types – customization, visualization, and social navigation (no example of personalization navigation was discovered). The digital museums that used customization, visualization, and social navigation were the Art Gallery of Ontario, Art UK, and the Smithsonian Learning Lab. The sites that included visualization and social navigation were the National Portrait Gallery, Cleveland Museum of Art, and SHOW.ME. The digital museums that contained customization and social navigation were WikiArt and Discover Islamic Art. The Web Gallery of Art only contained visualization navigation. The most common advanced navigation feature was social navigation, which allowed users to navigation by or add their own information to the content (tags). The second most common was visualization navigation. For example, a few of the digital museums used maps for to allow users to see the geographical information connected to artworks, either the venues (Art UK) or where the artist painted them (Web Gallery of Art). Customization was found in five of the nine digital museums. It was available in the digital museums that users could create an account for (allowing them to make albums of artwork for example).

Search systems were far more complex to compare and contrast. The digital museums had multiple search systems (the Cleveland Museum of Art had nine search systems alone). However, the primary search systems will be taken into account when comparing them (luckily the sites that had multiple systems were generally similar). The recall and precision for the sites varied between systems. There were six out of the eleven digital museums that prioritized precision. For example, a search in WikiArt only returned results that contained the query (in the title or description). The five other search systems prioritized recall over precision. For example, the Discover Islamic Art digital museum had the least amount of items in its catalogue, so recall returned more items for the users (helped by the synonym ring, the only one discovered). Creating criteria for this component of the search system entirely depended on the museum examined (and how the site had structured the rest of the search system).

The majority of the search interfaces were very simple, a search box with a magnify glass icon (Art Gallery of Ontario, National Portrait Gallery, Cleveland Museum of Art, WikiArt, SHOW.ME, and Smithsonian Learning Lab). Four of the virtual museums had complex search interfaces. These included the ones that had multiple filters and multiple search boxes. For example, even the simple search interface in Web Gallery of Art had two search boxes (one for artist and one for text) and two drop-down filters (Time-line and Form). This interface also had two buttons ("Search" and "Clear"). It was important to note that even though the Smithsonian learning lab had multiple search interfaces most of them were simple.

Advanced search interfaces (or features) were present in six of the nine digital museums. Only the Smithsonian Learning Lab, WikiArt and SHOW.ME learning digital museums did not have this feature. For Art UK and Discover Islamic Art the advanced search functionalities were present as the main search interfaces (filters and date limiters for example). The Art Gallery of Ontario (in the library and archive subsite), National Portrait Gallery, Cleveland Museum of Art (in the library archive search subsite) and Web Gallery of Art all had separate advanced search interfaces. For example, the Web Gallery of Art had an advanced search interface (accessed beneath the "Search" top-level category) with multiple boxes and filters.

The results displayed depended not only on the digital museum, but also the content returned. For example, the Cleveland Museum of Art returns 128 results per page, while Discover Islamic Art displayed ten per page. The Cleveland Museum of Art was able to display this many items because it displayed smaller images organized in a thumbnail/grid pattern (metadata was only shown when users hovered their cursor over an image). Most of the results were displayed like Discover Islamic Art, in a list format with 10 to 24 per page. The information returned was similar between the nine digital museums. This included an image, title, artist, date and description. Discover Islamic Art includes the location of an item, but did not have a description (this was true for Art UK as well). In SHOW.ME, the main search returned results that contained the "type" of object (collection, game, website, etc.). The digital museum that returned the most information was the Web Gallery of Art. This included image, artist, title, date, medium, size, collection, link to other works, image size, file colour, and file size. The information returned depended on the metadata that the content had and what information the sites thought their users wanted.

The main ranking/sorting method was relevance. This needed to be determined by testing the search system; the digital museums did not usually state that this method was used (though Art UK did). All the digital museums except the Web Gallery of Art ranked the results (at least initially) by relevance. The Web Gallery of Art sorted the search results alphabetically (by artist name, ascending). Four of the digital museums provided sorting options (the Cleveland Museum of Art, National Portrait Gallery, Art UK, and Smithsonian Learning Lab). These options included sorting the results alphabetically (ascending or descending) and chronologically (newest to oldest and vice versa).

The most popular additional action was an easily accessible search interface (sometimes populated with the search query) to conduct a new search. All nine of the digital museums had this option (by providing a search interface at the top of the results or elsewhere on the page). The digital museums that had other options included National Portrait Gallery (narrow/refine search), Cleveland Museum of Art (narrow/refine search and save search), Art UK (narrow/refine search, save a subset of results, and save search), WikiArt (save a subset of results), Discover Islamic Art (narrow/refine search and save a subset of results), and

Smithsonian Learning Lab (narrow/refine search and save a subset of results). The second most popular additional action was narrow/refine search (five of the digital museums had it). The third was saving a subset of results (four of the digital museum). Finally two of the digital museums let users save the whole search.

The most common query builder was autocomplete/autosuggest. There were two digital museums that had the full robust version, Art UK and WikiArt. Those two digital museums provided dropdown menus with preset queries (users could select artwork title, artists, venues, etc. all formatted according to the sites authority files). The rest of the digital museums (except the Smithsonian Learning Lab) had weaker autocomplete/autosuggest features. When users began to type in the search box, past queries were presented in a dropdown list (if they started with the letter(s) the user began to type). Five of the nine digital museums had some stemming capabilities (e.g., a search for paintings returned results with the term paint). Only one of the digital museums had spellchecking (the AGO main page search) and only one had natural language processing tools (the Smithsonian Learning Lab, though adding the question mark reduced the number of search results).

Only two digital museums did not have any advanced query languages supported in its system (WikiArt and SHOW.ME). Most of the search systems supported Boolean languages (and, or, and not) – the Art Gallery of Ontario, National Portrait Gallery, Cleveland Museum of Art, Art UK, Discover Islamic Art, Smithsonian Learning Lab, and the Web Gallery of Art. Four of the digital museums also striped out the stop words (a, the, is, at, etc.) in a query (Art Gallery of Ontario, Cleveland Museum of Art, Smithsonian Learning Lab, and Web Gallery of Art). For example, Discover Islamic Art did not strip out stop words because when they were added to a query the number of results went down (e.g., mosaics vs. the mosaics).

The content indexed for searching was very similar for the nine digital museums. Most of the digital museums indexed the same content – metadata (including indexing by topic), destination pages, full-text indexing (common for returning blog posts or stories), and by recent content (for events, exhibitions, and creation of content). The Web gallery of Art indexed the metadata (though this included descriptions) and destination pages (in this case only artwork).

All the digital museums had examples of descriptive, structural, administrative, and embedded metadata. The three types of digital museums had similarities in their metadata. The most common descriptive metadata values were title, description and keywords. There were a few digital museums (the Art Gallery of Ontario, Discover Islamic Art, and Web Gallery of Art) that only had title and description descriptive metadata. Other digital museums had much more. For example, the Smithsonian Learning Lab had notes (title and category) and SHOW.ME had additional teacher tags and topics. The Web Gallery of Art, Art UK and WikiArt all had artist pages with their own metadata. The descriptive metadata for artists included nationality, art movement, school, and description.

The structural metadata for the nine digital museums used different terminology, but they described many of the same values. The brochure museums all had size, medium/material, and type/format. The content museums all had medium and measurements. The learning museums had medium, physical description, and type. The similarities could be seen when comparing the structural metadata between the digital museums. For example, medium was found in all three. Size, measurements, and physical description all contained the same or similar content. Finally both brochure museums and learning museums had the type/format value (this defined what they object was). Some of the sites had much more structural metadata (Art UK, Web Gallery of Art, and Smithsonian Learning Lab). For example, the Art UK (in the shop) had the additional structural metadata for print size, frame size, print paper, etc.

The administrative metadata not only differed between the types of digital museums but also between each of the digital museums (this metadata depended on the business). However, there were some similarities, all the digital museums had artist name and date in the administrative metadata. The brochure digital museums also had call number and provenance metadata. The content museums had the additional fields – location, period, provenance, license, and photo credit. WikiArt also had (for artists) born, died, active years and URLs. It was learning digital museums that only had artist/creator name and date in common, because the administrative metadata was specific to each organization. For example, SHOW.ME also had held at, and production place (though this could be classified as provenance). The Smithsonian

Learning Lab included identifier, view original, additional information, and notes (contained within and contact information). Finally the Web Gallery of Art had location, artist birth/death dates (in addition to the artwork date) and catalogue number. Administrative metadata generally includes the artist/creator name, dates, history, and item/catalogue numbers.

Embedded metadata existed in all the digital museums, but some of them had almost none, while others had a lot. The most common embedded metadata was title, description and keywords. However, it should be noted that that metadata could apply to the overall site (for outside search systems) or for the particular objects. Other embedded metadata included open graph (in all the digital museums but Discover Islamic Art, Web Gallery of Art, and SHOW.ME), Dublin core (only Cleveland Museum of Art had this), and meta: itemprop fields (the Smithsonian Learning Lab).

All of the digital museums had some component of a controlled vocabulary. While evaluating the select digital museums, it was tested to see if they had synonym rings (only one), authority files, classification schemes, and/or a thesaurus. The hardest digital museum to evaluate was the Art Gallery of Ontario, but even it had authority files in a subsite (the Malcolmson Collection). The only digital museum with a discernable synonym ring was Discover Islamic Art. For example, a search for “Vase” returned “Pots” and “Vessels” (and the metadata of those items did not contain mentions of the word “Vase”).

Most of the digital museums had authority files for the artist names. The only two that did not were the Art Gallery of Ontario (it had one in a subsite, but it was not site wide) and SHOW.ME. The Smithsonian Learning Lab was the only site that used a pre-existing authority file standard, the Library of Congress Authority Files. The rest of the digital museums used site-specific authority files (that occasionally had similarities with standards, but did not follow them exactly). For example, in WikiArt it was “Raphael (1483-1520),” but in the Library of Congress it was “Raphaël, 1483-1520.”

All of the digital museums had classification schemes, found in the filters and facets users could use to search or browse. Even the AGO had a classification scheme for the events (at the physical location); users could narrow results using categories (exhibitions, food & drink, screenings, talks, etc.). The most common classification schemes found in the digital museums were type (all but the Discover Islamic Art), art movement (Art UK, WikiArt, Smithsonian Learning, Lab, and Web Gallery of Art had it), and place/location (Art Gallery of Ontario, National Portrait Gallery, Art UK, and WikiArt had it). Only one digital museum used a standard controlled vocabulary – the Smithsonian Learning Lab (using the Library of Congress Subject Headings). The classification schemes used by the digital museums depended on the metadata and content of the site. However, they all had at least two classification schemes connected to the content (in order to facilitate searching and browsing).

All the digital museums had collections of predefined terms used to connect content together (as seen in the classification schemes above). These terms could be structured in hierarchy relationships (parent-child) to show what the digital museums thesaurus could look like. It was important to note that most of the thesaurus information came from an interpretation of the sites index terms (metadata, keywords, etc.), in order to show the reader the preferred terms. Only digital museums that followed a pre-existing controlled vocabulary (the Smithsonian Learning Lab, which used the Library of Congress Subject Headings) were definite. Lacking keywords/rich classification schemes hindered the identification of the sites thesaurus, as was the case in the Art Gallery of Ontario. The other seven digital museums had enough keywords and classification schemes managed by the site (not added by users, which would likely not follow preferred terms) to decipher the thesaurus used (the National Portrait Gallery, Cleveland Museum of Art, Art UK, WikiArt, Discover Islamic Art, SHOW.ME and Web Gallery of Art). These thesauri supported both searching and browsing, which was determined by how they were discovered (through the filters and/or facets of the site).

All the digital museums had semantic relationships between terms used on the sites. These could be found in the hierarchy relationships of the constructed thesauri or the associative relationships between classification schemes (both between the subcategories and how the classification schemes were used as filters within a site). Equivalence relationships were much harder to determine, they could only be seen in the Discover Islamic Art (through the glossary) and the Library of Congress Subject Headings (which the Smithsonian Learning Lab used). For example, the glossary in the Discover Islamic Art presents users with synonyms and alternate spellings for the term users looked up.

Not all of the digital museums had a rich fully functioning faceted classification. For example, in Discover Islamic Art users could browse using facets, but only one at a time (e.g., period/dynasty). WikiArt had many different facets, but users could only select one at a time beneath the categories “artist” or “artworks” (e.g., looking at artists by nationality). In SHOW.ME the faceted classification system allowed users to access content by theme (or select types), but then they could be narrowed/sorted by type as well. The Art Gallery of Ontario used faceted classification in one of the subsite collections (the Boxwood Collection) but not in the main site. The rest of the digital museums (National Portrait Gallery, Cleveland Museum of Art, Art UK, Smithsonian Learning Lab, and The Web Gallery of Art) all had more traditional faceted classification. These digital museums provided users with the opportunity to browse or search (or narrow search results) using the classification schemes discussed above. For example, in the Web Gallery of Art users could browse the artist index using the facets – school, period, time-line, and profession (all with predefined terms).

Summary

Here are the information architecture criteria as determined by the heuristic evaluation:

- All the digital museums used hybrid organization schemes (commonly, ambiguous organization schemes for the top-level categories and exact organization schemes for organizing content).
- All the digital museums used hybrid organization structure (with a polyhierarchical top structure and database oriented bottom structure). Focused entry points were used by the content and learning digital museums.
- All types of labels were used (contextual link labels, headings, navigation labels, index terms and icon labels). Digital museums need to keep labeling consistency in mind.
- All the digital museums used global, local, contextual/hypertext, and utility navigation. Five of the nine digital museums used breadcrumb navigation.

- Supplemental navigation depended on the digital museums and the content, but it was advisable to use indexes, guides, control panels and pagination navigation. More than half of the digital museums had those features.
- The advanced navigation features found were customization, visualization, and social navigation. If the digital museum had user accounts they should offer customization and social navigation. Visualization navigation worked well to help contextualize where the content came from.
- Search systems depended on the content and the digital museums.
    - Precision was used more often than recall (slightly – six out of the eleven examined).
    - Most of the search systems used a very simple search interface (box with a magnify glass icon button). Four search systems had complex interfaces (to counteract the lack of filters/facets available to narrow down the search results).
    - Advanced search functionality should be available (six of the nine digital museums offered this feature).
    - Results displayed should be in a list/grid format with a lower number per page (10 to 24). The information that should accompany them includes image, title, artist, date and description.
    - The search results should be sorted by relevance (eight of nine). Sorting options (four out of nine had this option) would be a good idea as well, alphabetically (A-Z) and/or chronologically (date).
    - The most popular additional action was adding a search interface near the results so that users could easily conduct a new one. Adding filters to narrow/refine the search results (five out of nine had this feature) and saving a subset of results (four out of nine) are suggested as well.
    - Select digital museums should have autocomplete/autosuggest query builders (either the robust kind or the type that just saved the uses search history).
    - The search system should support Boolean languages (seven out of nine) and could strip out stop words (four of nine).
    - The metadata (indexing by topic), destination pages, full-text indexing (for blog posts, etc.), and recent content should all be indexed for searching.

- All selected digital museums should use descriptive, structural, administrative, and embedded metadata.
- All selected digital museums should use authority files (especially for artist names).
- All selected digital museums should have classification schemes, which users could use to filter content with.
- All selected digital museums should have some sort of thesaurus (either following a pre-existing example or a thesaurus standard). With semantic relationships (hierarchy, associative and equivalence) that connected the preferred terms.
- All select digital museums should have faceted classification (used particularly for searching and browsing).

Bugs/Issues

There were eighteen bugs/issues found in the digital museums evaluated. Fixes for the bugs ranged from simple to complex. The broken links found in the Art Gallery of Ontario (to Flickr), Cleveland Museum of Art (#paradethecircle empty page and CMA Tumblr page under construction), SHOW.ME (link to a dinosaur dig in a different museum), and Web Gallery of Art (page not found for "Director of Online Museum" link) could be fixed or deleted. If the page it linked to no longer existed then this link should be deleted (e.g. the SHOW.ME dinosaur link). If the page moved then the link should be updated.

The first bug found in the Art Gallery of Ontario was the disorganization of an alphabetical list (in the spoliation research projects, there was a "D" entry above "B"). The digital museum needed to edit their HTML, so that the entry appeared in the correct order. The floating categories found when the user navigated to the "Group Visits" page (in the local navigation) should be accessible from other pages (preferably the top-level category "Visit" that the breadcrumb navigation on that page indicated it was organized under) or deleted (if the site did not want users to access that information then it should not be there).

In the National Portrait Gallery the advanced search system had a drop-down menu that extended well off the bordered section dedicated to that function. The drop down menu was for “Portrait set” and the option creating this issue was “Portraits of Member of the Society of Painters in Water Colours 1864: photographs by Cundall, Downes & Co and other, 1850s to 1860s.” The dropdown menu extended its size to the longest option it contained. To fix this issue, shorten the title using “…”, like many of the other options available in that dropdown menu did (e.g., “Political sketches by H.B….”). When the search system in the National Portrait Gallery returned many results (under “Other Pages”) the pagination navigation extended off the page and across local navigation (because it listed all the pages). To fix this the National Portrait Gallery should use arrows (like other sections of that search system used). For example, the results would look like this: < 1, 2, 3, 4, 5 …. 64 > rather than 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.

The Cleveland Museum of Art had some CSS issues. When in the “learn” top-level category, the menu that popped up in the lower left corner did not come up the whole way (it cut off the site map information). When searching, the “clear” button (used to clear the search and start again presumably) was cut off, so the user could only see the “Cle” part. This could be fixed using the CSS overlay option or overflow option (discussed in more detail below). If the window was too small (around 7 inches wide) the utility navigation text overlapped. These were all CSS issues that the site needed to deal with. For example, to fix the overlapping information in the footer, designers could change the CSS, so the height of the footer increased or decreased depending on the screen size (changing the explicit height field) (Kershaw, 2013).

In Art UK, the button that sent users back to the top of the pages didn’t work on the homepage (but it did on other pages in the site). To fix this, the website designers need to adjust the coding (JavaScript for example). Also in Art UK, when in the shop pages, the organization of the global navigation menu changed (moving the shop information from the bottom to the top of the menu). This may not be an issue, the site could have done it on purpose, but if it was an error the designers would need to change the HTML code (so that the shop information remains last on the main menu list).

There was a CSS issue found in the glossary of the Discover Islamic Art digital museum. The glossary did not let users scroll through the list of terms (they could use the up or down arrows (on their keyboard) to select the terms off the page, but that still did not move the list down). To fix this, the site designers need to adjust (or add) the CSS overflow option. This adds a scrollbar to the content if it is too big/long to fit in a specified area (w3schools.com, n.d.). This feature was available in the glossary definitions (shown when users selected a term to view). In WikiArt, if the screen size of the browser was too small the control panel fields (under "quick edits") broke and overlapped. This would be a matter of adjusting the CSS of the site (overflow, see above). The final bug that needs to be discussed was also a CSS issue (with overflow again). In the Smithsonian Learning Lab (under the user dashboard) the information within the collections that users had added to a list of favorites overlapped. To fix this the site needed to either make the description fields bigger or provide the CSS overlap feature.

Chapter Summary

The data collection and analysis for this thesis was conducted in two stages. The first part, a content analysis study, began with a pilot study in order to create the subcategories (organized beneath organization, navigation, labeling, search, and vocabulary systems) and improve the coding frame (see Appendix B and G). The main study used the improved coding frame and created subcategories to examine four general knowledge information architecture books (including the source used in the pilot study). Once data collection was completed, the research was examined to view the relationships both between the sources and between the categories. For example, the source used in the pilot study had the most coded entries and the fourth source had the least. The heuristics were created by reviewing the coded text, defining entries as a definition, example, pro or con, and component (see Appendix K). These were compiled into a list and were used to inform a guideline document that would be used in conjuncture with the definitions during the heuristic evaluation (see Appendix L). For example, how to find and explain labeling consistency issues (and provide good and bad examples).

The heuristic evaluation examined nine select digital museums, three of each type (brochure, content, and learning). The brochure digital museums were the Art Gallery of Ontario, National Portrait Gallery, and Cleveland Museum of Art. The content digital museums included Art UK, WikiArt, and Discover Islamic Art. The learning digital museums were the Smithsonian Learning Lab, SHOW.ME, and Web Gallery of Art. To become familiar with the site's audience, content, and context the contextual framework for each digital museum was defined. For example, many of the digital museums could list researchers as a main audience type, which you can see by the way that the websites labeled different sections of their site (e.g., "Research"). The heuristic evaluation examined the nine digital museums with the heuristic definitions and guidelines (see Appendix K and L). Twenty-eight or more hours were spent evaluating the information architecture for each of the digital museums, identifying the heuristics found in the site as well as any bugs discovered. When the heuristic evaluation was completed the information architecture components were compared between the different types and then all together (the latter found in the findings section). This created the final list of information architecture criteria. For example, hybrid organization structures and schemes were criteria that all of the evaluated digital museums used. The hybrid categories combined different IA principles to structure and organize information, allowing a flexible presentation of the content for users. An additional advantage of heuristic evaluations was finding bugs/issues associated with IA in the evaluated digital museums. These were retested (to exclude any false positives) and solutions suggested. The solutions mainly revolved around deleting or replacing broken links and fixing the CSS.

Chapter 5: Conclusion

Introduction

This research study created a list of criteria that select digital museums could utilize when designing the information architecture of their website. This will facilitate consistency between cultural heritage websites and provide more effective access to information for users. When users know how to use a site and can easily understand the information presented on them, they gain confidence in the website and themselves (and are much more likely to return) (Parandjuk, 2010; Simon, 2008; Spencer, 2011; Srinivasan et al., 2009; Wodtke & Govella, 2009). To accomplish the creation of the criteria, two research methods were completed. This study examined nine select digital museums in a heuristic evaluation in order to determine the information architecture principles that they use in the design of their site. The heuristic evaluation was based on the heuristics created during a formal content analysis study.

Project Summary

The research study began with a formal content analysis study that examined four general knowledge information architecture books. After selecting the sample, a coding frame was created to guide the study. This included coding instructions, relevant and irrelevant materials, and definitions of the main categories the text was coded into (organization, navigation, labeling, search, and vocabulary systems). This was tested in a pilot study (with *Information Architecture for the Web and Beyond*). After completing the pilot study, new subcategories were organized beneath the main categories as well as changes made to the coding frame. With the final coding frame established, the main study was conducted with the four information architecture books.

Data collection for the main study of the content analysis went through each chapter recording and coding text that was directly related to information architecture principles. It was during data analysis that the list of heuristics was created for the next stage of the study. To do this the data was organized by the codes so that all the coded text was grouped by the categories and subcategories. Any information in the miscellaneous categories (residual categories to catch information that was not found in the pilot study source) was examined and then folded into the final list of heuristics. The final list of heuristics contained 119 principles. From these principles a guideline of questions were created to act in conjunction with the list of heuristics for the heuristic evaluation.

The heuristic evaluation examined nine select digital museums (three of each type brochure, content, and learning) – the Art Gallery of Ontario, the National Portrait Gallery, The Cleveland Museum of Art, Art UK, WikiArt, Discover Islamic Art, Smithsonian Learning Lab, SHOW.ME, and Web gallery of Art. In this study there was only one evaluator, who was an expert in both usability and the domain (work experience as a usability coordinator and writing help documentation, with an undergraduate degree in Art History), who examined the nine select digital museums three times. To become familiar with the sites, a contextual framework was created for each digital museum (defining the users, content, and context for the site). Then the heuristic evaluation began, looking for both bugs and information architecture principles digital museums use in their live websites. The list of bugs was recorded in a Microsoft Excel spreadsheet and the information architecture principles were noted down in report format (see Appendix M for an example). The third examination happened at the end of data collection and the beginning of data analysis, to confirm the results of the heuristic evaluation and examine the bugs found to rule out any false positives (there ended up being four false positives). Once the evaluations all the nine digital museums were completed data analysis began.

To create the list of criteria, the digital museums information architecture principles were compared and contrasted – first between the different types (brochure, content, and learning) and then between all nine of the digital museums. For example, the presence of breadcrumb navigation features was discussed in the content digital museums (WikiArt and Discover Islamic Art did have this feature, ART UK did not) as well as where in the site it could be found. This feature was then compared between all nine of the digital museums. Using the second comparisons the criteria were created, if all of the nine digital museums contained the same information architecture principle it was highly suggested that this be used in the design of select digital museums (e.g. they all used hybrid organization schemes and structures). The number of digital museums that had the IA principles was noted so that the reader could choose if they wanted to follow it (e.g., five out of nine digital museums had breadcrumb navigation). The final list of criteria can be seen on page 165. Further research could expand this study and include volunteer participants to find out if they agree with the criteria assembled in this study.

The data collection and analysis chapter closed with an examination of the bugs/issues found during the study and suggestions for possible solutions. This could be as simple as deleting or adding the correct URL to fix broken links. Or as complex as fixing overflow problems by working with the CSS of a site (e.g., in Discover Islamic Art the glossary was not scrollable, so you could not see all the option under each letter). For the full list of suggestion see page 167.

Data collection and analysis was presented to the reader in rich detail so that they can understand the decisions made throughout the study and draw their own conclusions (which increases the reliability of the research). This helps establish the internal validity of the study, as did the amount of time spent on the heuristic evaluation (on average each digital museum was evaluated for twenty-eight hours, going over the interface three times) and the use of quantitative methods (statistics) during the content analysis study. The external validity was established through the design of the methodology (other researchers could use it to confirm the results), as well as the detailed description in the collection of the data (the bug spreadsheet and the report form for the information architecture principles).

Contributions and Implications of Research

The criteria created in this study can guide the design and implementation of information architecture in select digital museums, improving consistency and interoperability between websites (Riley-Huff, 2012; Rosenfeld et al., 2015; Simon, 2008; Teather, 2008). This means that the information architecture between select digital museums will have the same/similar IA organization, labeling, navigation, search, and vocabulary systems. These similarities not only improve consistency and interoperability between websites, but also the usability, findability, and understandability of a website for users (key IA concepts) (Rosenfeld et al., 2015; Spencer, 2011; Wodtke & Govella, 2009). For example, when encountering a new website, if the IA foundation is recognizable (e.g., how users navigate to the visitor information or how the search results are presented) a sense of familiarity between the user and the website is created. Familiarity in a website means that the site is easier for them to use (usability), understand (understandability), and find information within (findability) – increasing the confidence that users have in both the website and themselves (Parandjuk, 2010; Simon, 2008; Spencer, 2011; Srinivasan et al., 2009; Wodtke & Govella, 2009). The criteria listed in this study can also be used to improve or guide digital museums with poorly designed information architecture (which would need to be identified using further usability studies). This would be highly helpful as poorly designed IA can decrease revenue and the number of people who visit a website (Toub, 2000; Wodtke & Govella, 2009). This list of criteria can be helpful for select digital museums looking to improve how users access, find, and understand the information in their site.

The criteria can only increase consistency and interoperability between the computer interfaces of digital museums and not across platforms. Though the criteria in this study may be suited for different platforms (smartphone, tablets, etc.) those interfaces were not evaluated. It would be outside the scope of the study to suggest that the different platforms could be improved upon by using the criteria. This will be discussed in further detail below.

How the sample was selected meant that the findings of this study were not generalizable to all digital museums because a probability sample could not be selected (e.g., each digital museum in existence was selected randomly, this number is not known) (Bickman & Rog, 2009; Leedy & Ormrod, 2010; Merriam & Tisdell, 2016). The best that this study can do (to ensure the validity and reliability of the research) is applying the criteria to digital museums similar to those evaluated in this paper (Merriam & Tisdell, 2016). Additionally, by evaluating different types of digital museums (brochure, content, and learning) it means that there is a variation of the sample and so can be applied to more digital museums than if there was just one type studied. To not reach beyond the scope of the study the criteria cannot be applied to digital museums that were excluded from the sample (e.g., large digital museums like Artstor and The Metropolitan Museum or Art, non-English digital museums, etc.). However, it could be possible to apply the criteria to digital museums that have not been recently updated, because they are improving their IA to the level of the sites that were evaluated. Additionally, the contextual framework should be considered when applying the list of criteria to a digital museum, the sites should have similar audience and content for the best results. Select smaller websites and currently developing digital museum websites could use this list to make sure that their IA design is comparable to other digital museum sites (their graphic design and content can obviously vary). Here is an example of websites that could use the criteria to improve their IA design or promote consistency and interoperability between sites.

- The Seattle Art Museum (brochure)
    - http://www.seattleartmuseum.org/visit
- The Glenbow Museum (brochure)
    - http://www.glenbow.org
- The McManus, Dundees Art Gallery & Museum (brochure)
    - http://www.themcmanus-dundee.gov.uk
- Museum Crush (content)
    - http://museumcrush.org
- Tyne & Wear Archives & Museums (content)
    - https://twmuseums.org.uk
- Sharing History (content)
    - http://www.sharinghistory.org
- International Museum of Women (learning)
    - http://exhibitions.globalfundforwomen.org
- World Images Kiosk (learning)
    - http://worldimages.sjsu.edu/?sid=14729&x=440604
- Virtual Museum of Canada (learning – studied in the 2012 pilot study, but the website has been redesigned)
    - http://www.virtualmuseum.ca/home/

It should be noted that if the criteria is applied to those digital museum, it should only be done in similar context (to not exceed the scope of the study). Additionally, these digital museums represent possible websites that the criteria can be applied to, they may not need the criteria presented in this study or they could already follow aspects of the criteria (as it was created through the comparison of live websites).

Areas for Future Research

There are many areas for future research based on the results of this study. This included using different research methods to conduct the same study and/or expanding the study to include user participation. For example, a future study could build on these results and include heuristic evaluations conducted by volunteer participants on the same digital museums to see if they come to the same conclusions. Do they find the same information architecture components? Do they find the same bugs? In all likelihood more evaluators would find more problems with the digital museums interfaces and produce findings that (if there was agreement between evaluators) could be standardized. No matter what the results, the amount of time or the number of digital museums would have to be reduced. Volunteer participants are unlikely spend three weeks evaluating the nine digital museums. Additional ways that the heuristic evaluation could be expanded include adding more museums, studying larger more complex digital museums (like Artstor), or evaluating the mobile/tablet interfaces of the digital museums.

The content analysis portion of this study could also be expanded to include additional coders and/or additional information architecture sources. Multiples coders can test the validity and reliability of the coding frame by calculating the coefficient (the agreements between coders). The results of this study could then be used to evaluate the same digital museums to see if the new heuristics (if they changed at all) had any effect on the results of the study. Alternately, the heuristics created from this content analysis study could be used for different studies, to evaluate different websites information architecture (e.g., digital libraries, archives, etc.).

Using the information criteria created in this study, a prototype digital museum (digital or wireframes) could be developed. Users could test the prototype (in a guided evaluation) to see if they agreed with the criteria presented in this study or if changes would be necessary.

Ideas for further research could be taking the concept of this study (establishing information architecture criteria for digital museums to promote consistency and access to content) and use different methodologies to see if the same criteria would be produced. A user study could be conducted on digital museums visitors, to see why they visited a website (though this in itself would be a full research study). The methods used to find out this information include online surveys, questionnaires, focus groups, web analytics (viewing where they went, when they visited a site, and how long they spend completing tasks on the digital museum websites), etc. Once the types of museum visitors are defined, personas and scenarios could be constructed, used in either a cognitive walkthrough (evaluators navigate through websites according to specific tasks that the personas would complete) or heuristic walkthrough (traveling through a digital museum using heuristics as the created personas). In the later usability inspection method, the heuristics created in this study could be used. It would be interesting to see if the criteria presented in this study stayed the same, or if evaluating a digital museum using personas of their target audience would change the results.

Another area of study is updating the definitions for the different types of digital museums. As technology advances the brochure, content, and learning digital museum definitions have started to blur together (e.g., many content museums provide rich contextual links to educational information or brochure digital museums are now so large and complex that they could be considered an amalgamation of all three types). More types should be added to the list, for example, what about digital museums that only have born digital collections. Would they be content or brochure digital museums? A content analysis or a more informal survey of digital museums on the web could create new classifications of digital museums.

Final Thoughts

Digital museums and information architecture are two areas of study that I am very interested in. Providing access to cultural heritages objects (paintings, sculpture, digitally created object, tapestry, etc.) is personal to me because I have spent a lot of my time on these sites over the years. This is why I think organizing the information architecture of digital museums is so important, I have been the user that has become frustrated with a site when I cannot find the information I am looking for.

I planned this study to be as valid and reliable as possible because access to information needs to be a priority for these sites. This was achieved mainly through rich description when presenting the methodology design, data collection, data analysis, and findings of the study. I chose to do a content analysis study because I wanted to create the heuristics using a formal method that let the reader see the decisions I made. The usability inspection method, heuristic evaluation, was originally suggested by one of my advisors, but once I began learning about the method I knew it was the best choice for this study. The heuristic evaluation allowed me to immerse myself in a websites, focusing on how digital museums use information architecture.

Some of the digital museums examined for this thesis were far more complicated than others. I began this portion of the study evaluating the Art Gallery of Ontario (AGO), incorrectly thinking that this would be a less complicated site (it still surprises me how wrong I was about this). The AGO was by far the most complex and poorly IA designed digital museums in this study (and would benefit from IA criteria). The switching between different "homepage" designs, which had different labels, navigation, and search systems made evaluating this site very difficult. It took me two weeks to evaluate this digital museum (working approximately 8-10 hours a day); about the same amount of time it took me to evaluate the rest of the sites. This is partially due to the poor design and partially because it was the first digital museum examined.

All the brochure digital museums were complex. The content on those sites not only included information about the physical location but as well as collection of objects and contextual information (blurring the lines between the types of digital museums). Comparably the content digital museums information architecture was straightforward and easily identifiable. All three of the sites were fun to use and had fascinating content. The learning digital museums were interesting, especially the Smithsonian Learning Lab. This digital museum let users create their own collection, they could add new objects and make connections between existing content in order to create learning resources. This was not something I had encountered before and I enjoyed evaluating it.

The list of criteria created by comparing the information architecture found in the evaluated digital museums involved a lot of description and examples. It is hard to describe how heading labels are used in a site without describing their location and giving examples. This resulted in rich detail, but a long chapter. The final list of criteria was presented in the findings section of Chapter 4. I believe that these criteria are sound, but they would benefit from further research (to triangulate the results).

The bugs/issues were found throughout the evaluations. I tried really hard to make sure that there were no false positives, but on reexamination there ended up being four. There were not that many bugs/issues, but I did not expect there to be since I was examining live websites. I researched solutions for each of the bugs found. I was surprised at how many of the issues were related to CSS, considering that they did belong to IA heuristics. This showed how interconnected information architecture is in an interface.

The methods used in this study provided a way to both define the information architecture components and conduct an evaluation of digital museums using those principles. The list of criteria created from the methodology includes recommendations for each component found, organized beneath the IA systems. The bugs/issues discovered over the course of the study were reevaluated and solutions suggested to improve the IA in the live websites. All digital museums have information architecture, but the design of the IA components varies between sites and some are poorly implemented. Creating a list of information architecture criteria will not only improve current IA practices in select digital museums but also promote interoperability between sites and consistency in how users access information.

References

ABC-CLIO (2013). Navigation. Retrieved from: http://www.abc-clio.com/ODLIS/odlis_n.aspx

ABC-CLIO (2013). Search. Retrieved from: http://www.abc-clio.com/ODLIS/odlis_s.aspx

Aiello, L. C. (2016). The multifaceted impact of Ada Lovelace in the digital age. *Artificial Intelligence*, 235, 58-62. https://doi.org/10.1016/j.artint.2016.02.003

Alexander, E. P., Alexander, M., & Decker, J. (2017). *Museums in motion: An introduction to the history and functions of museums*. Lanham, MD: Rowman & Littlefield.

Amsel-Arieli, M. (2012). Cabinets of curiosity (wunderkammers). *History Magazine*, 13, 40-2. Retrieved from: http://blogs.iac.gatech.edu/bellyful/files/2012/11/Cabinets-of-Curiosity.pdf

Andre, T.S., Hartson, H.R., & Williges, R.C. (2003). Criteria for evaluating usability evaluation methods. *International Journal of Human Computer Interaction*, 15(1), 145-181. http://dx.doi.org/10.1207/S15327590IJHC1304_03

Anganes, A., Pfaff, M. S., Drury, J. L., O'Toole, C. M. (2016) The heuristic quality scale. *Interacting with Computers*, 28 (5), 584-597. doi: 10.1093/iwc/iwv031

Antonaci, A., & Ott, M. (2014). Virtual museums and learning innovation. *The International Scientific Conference eLearning and Software for Education*, Vol. 2, 28-33. https://doi.org/10.1093/iwc/iwv031

Art Gallery of Ontario (2017). Our mandate. Retrieved from: http://www.ago.net/mandate Accessed May 28, 2017.

Art UK (2017). Our mission. Retrieved from: https://artuk.org/about/our-mission Accessed June 10, 2017.

Art UK (2017). Welcome. Retrieved from: https://artuk.org/about/welcome Accessed June 10, 2017.

Babbage, C. (2011). Passages from the Life of a Philosopher. Cambridge University Press. (Original work published 1864). Retrieved from: https://books.google.ca/books?id=Fa1JAAAAMAAJ&oe=UTF-8&redir_esc=y

Baca, M. (2003). Practical issues in applying metadata schemas and controlled vocabularies to cultural heritage information. *Cataloging & classification quarterly*, 36(3-4), 47-55. http://dx.doi.org/10.1300/J104v36n03_05

Baca, M. (2004). Fear of authority? Authority control and thesaurus building for art and material culture information. *Cataloging & classification quarterly*, 38(3-4), 143-151. http://dx.doi.org/10.1300/J104v38n03_13

Barton, J. (2005). Digital libraries, virtual museums: Same difference?. *Library Review*. 54(3): pp. 149-154. https://doi.org/10.1108/00242530510588908

Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *Nursing Plus Open*, 2, 8-14. https://doi.org/10.1016/j.npls.2016.01.001

Benjamin, W. (2008). *The work of art in the age of mechanical reproduction*. (J. A. Underwood Trans.). London, UK: Penguin Books Ltd. (Original work published 1936).

Berners-Lee, T., Cailliau, R., Groff, J. F., & Pollermann, B. (2010). World-wide web: The information universe. *Internet Research*, 20(4), 461-471. https://doi.org/10.1108/10662241011059471

Bertacchini, E., & Morando, F. (2013). The future of museums in the digital age: New models for access to and use of digital collections. *International Journal of Arts Management*, 15(2), 60-72. Retrieved from: https://search.proquest.com/docview/1323008907?accountid=33837

R. Bianchini (2016, February 4). *When museums became virtual*. Retrieved from: https://www.inexhibit.com/case-studies/virtual-museums-part-1-the-origins/

Bickman, L., & Rog, D. J. (2009). Applied research design: A practical approach. In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (3-43). Thousand Oaks, CA: SAGE Publication, Inc.

Biedermann, B. (2017). 'Virtual museums' as digital collection complexes. A museological perspective using the example of Hans-Gross-Kriminalmuseum (University of Graz). *Museum Management and Curatorship*, 1-17. http://dx.doi.org/10.1080/09647775.2017.1322916

Bowen, J. P. (2010). A brief history of early museums online. *The Rutherford Journal*, 3. Retrieved from: http://rutherfordjournal.org/article030103.html

Bowen, J. P. (2017). Alan Turing: Founder of Computer Science. *Engineering Trustworthy Software Systems* (pp. 1-15). New York City, NY: Springer International Publishing. https://doi.org/10.1007/978-3-319-56841-6_1

Bowen, J. P. (2016, July). Alan Turing: virtuosity and visualisation. In Proceedings of the conference on Electronic Visualisation and the Arts (pp. 197-204). British Computer Society. http://dx.doi.org/10.14236/ewic/EVA2016.40

Bowen, J. (2000). The virtual museum. *Museum International*, 52(1), 4-7. doi: 10.1111/1468-0033.00236

Bowen, J. P. (1999). Time for renovations: A survey of museum web sites. *Museums and the Web 1999*. Retrieved from: http://www.museumsandtheweb.com/mw99/papers/bowen/bowen.html#JonesGarmil97ref

Bowen, J. P., Bennett, J. & Johnson, J. (1998). Virtual visits to virtual museums. *Museums and the Web 1998*. Retrieved from: http://www.museumsandtheweb.com/mw98/papers/bowen/bowen_paper.html

Bush, V. (1931). The differential analyzer. A new machine for solving differential equations. *Journal of the Franklin Institute*, 212(4), 447-488. https://doi.org/10.1016/S0016-0032(31)90616-9

Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176(1), 101-108. Retrieved from: https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/

Campbell-Kelly, M., Aspray, W., Ensmenger, N., & Yost, J. R. (2014). *Computer: A history of the information machine*. Boulder, Colorado: Westview Press.

Carreras, C., & Mancini, F. (2014). A Story of Great Expectations: Past and Present of Online/Virtual Exhibitions. *DESIDOC Journal of Library & Information Technology*, 34(2).

Carrozzino, Marcello and Massimo Bergamasco. (2010). Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage*, 11, 452-458. https://doi.org/10.1016/j.culher.2010.04.001

Chae, G., & Kim, J. (2011). Rethinking Museum Management by Exploring the Potential of Social Tagging Systems in Online Art Museums. *International Journal of the Inclusive Museum*, 3(3).

Cleveland Museum of Art (2017). Our mission. Retrieved from: http://www.clevelandart.org/about/history-and-mission/mission-statement Accessed June 10, 2017.

Cockton G., Woolrych A. (2001) Understanding inspection methods: Lessons from an assessment of heuristic evaluation. In: Blandford A., Vanderdonckt J., Gray P. (eds.) *People and Computers XV—Interaction without Frontiers*. London, UK: Springer.

Cody, S. A. (1997). Historical museums on the World Wide Web: An exploration and critical analysis. *The Public Historian*, 19(4), 29-53. Retrieved from http://www.jstor.org/stable/3379493 Accessed July 4, 2017.

Coe, I. R., & Ferworn, A. (2016). The life and contributions of countess Ada Lovelace: Unintended consequences of exclusion, prejudice, and stereotyping. *IEEE Technology and Society Magazine*, 35(4), 46-49. doi: 10.1109/MTS.2016.2618679

Cohen-Almagor, R. (2011). Internet history. *International Journal of Technoethics*, 2(2), 45-64.

Cohen, K., Elkins, J., Lavin, M. A., Macko, N., Schwartz, G., Siegfried, S. L., & Stafford, B. M. (1997). Digital culture and the practices of art and art history. *The Art Bulletin*, 79(2), 187-216. doi: 10.1080/00043079.1997.10786732

Connaway, L. S., & Radford, M. L. (2017). *Research methods in library and information science*, (6th ed). Santa Barbara, CA: Libraries Unlimited.

Conway, P. (2010). Preservation in the age of Google: Digitization, digital preservation, and dilemmas. *The Library Quarterly*, 80(1), 61-79.

Covert, A. (2014). *How to make sense of any mess: Information architecture for everybody*. SI: CreateSpace Independent Publishing.

Crawford, S. (2011, March). How information architecture works. *How Stuff Works: Tech*. Retrieved from http://computer.howstuffworks.com/information-architecture.htm/printable Accessed July 20, 2017.

Merriam Webster Dictionary (n.d.). Criterion [Def. 2]. Retrieved from: https://www.merriam-webster.com/dictionary/criterion

Ding, W., Lin, X., & Zarro, M. (2017). Information Architecture: The Design and Integration of Information Spaces. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 9(2), i-152. https://doi.org/10.2200/S00755ED2V01Y201701ICR056

Discover Islamic Art. (2017). About museums with no frontiers. Retrieved from: http://www.museumwnf.org/atrium_about.php Accessed June 10, 2017.

Duncan, C., & Wallach, A. (2004). The universal survey museum. *Museum studies: An anthology of contexts*, 51-80.

English Oxford Dictionary (n.d.). Principle [Def. 2.1]. Retrieved from: https://en.oxforddictionaries.com/definition/principle

Essinger, J. (2014). *Ada's algorithm: How Lord Byron's daughter Ada Lovelace launched the digital age*. Brooklyn, NY: Melville House.

Falk, J. H. (2016). *Identity and the museum visitor experience*. Abingdom, UK: Routledge.

Fantoni, S. F., Stein, R., & Bowman, G. (2012, April). Exploring the relationship between visitor motivation and engagement in online museum audiences. *Museums and the Web*.

Findlen, P. (1989). The museum: Its classical etymology and renaissance genealogy. *Journal of the History of Collections*, 1(1), 59-78. https://doi.org/10.1093/jhc/1.1.59

Fishkin, R, & Moz Staff (2015, December 18). Beginners guide to seo (search engine optimization. *Moz*. Retrieved from https://moz.com/beginners-guide-to-seo Accessed June 20, 2017.

Foo, Schubert. (2008). Online virtual exhibitions: Concepts and design considerations. *Journal of Library and Information Technology*. 28(4) 22-34.

Friess, E. (2015). Personas in Heuristic Evaluation: An Exploratory Study. *IEEE Transactions on Professional Communication*, 58(2), 176-191. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7108068&isnumber=7110677

Gray, W. D., & Salzman M. C. (1998). Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human–Computer Interaction*, 13(3), 203-261. http://dx.doi.org/10.1207/s15327051hci1303_2

Green, D. (2000). The evolution of web searching. *Online information review*, 24(2), 124-137. https://doi.org/10.1108/14684520010330283

Gould, J. D. & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, 28(3), 300-311. http://dx.doi.org/10.1145/3166.3170

Grice, G. (2015). *Cabinet of curiosities: Collecting and understanding the wonders of the natural world*. New York: Workman Publishing.

Hagen, H. A. (1876). The history of the origin and development of museums. *The American Naturalist*, 10(2), 80-89; 135-148. doi: 10.1086/271587

Harms, I. & Schweibenz, W. (2001). Evaluation the usability of a museum website. *Museum and the Web*, 2001. Retrieved from http://www.museumsandtheweb.com

The Metropolitan Museums of Art (2017). History of the museum. Retrieved from http://www.metmuseum.org/about-the-met/history Accessed July 10, 2017.

Hoptman, G. H. (1992). The virtual museum and related epistemological concerns. *Sociomedia, multimedia, hypermedia and the social construction of knowledge*. 141-159.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288. https://doi.org/10.1177/1049732305276687

The Information Architecture Institute (2017). What is information architecture? Retrieved from: http://www.iainstitute.org/what-is-ia Accessed April 25, 2017.

International Council of Museums (2007). Museum definition. Retrieved from http://icom.museum/the-vision/museum-definition/ Accessed June 29, 2017.

Jacobsen, N. E., Hertzum, M., & John, B. E. (1998). The Evaluator Effect in Usability Studies: Problem Detection and Severity Judgments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(19), 1336 – 1340. doi: 10.1177/154193129804201902

Karoulis, A., Sylaiou, S., & White, M. (2006). Usability evaluation of a virtual museum interface. *Informatica*, 17(3), 363-380.

Karp, Cary. (2004a). Digital heritage in digital museums. *Museum International*. 56(1-2): pp. 45-51. doi: 10.1111/j.1350-0775.2004.00457.x

Karp, Cary. (2004b). The legitimacy of the virtual museum. *International Museum Council of Museums News*. 57(3), 5. Retrieved from: http://www.cool.conservation-us.org/icom/fileadmin/user_upload/pdf/ICOM_News/2004-3/ENG/p8_2004-3.pdf

Kaushik, A. (2007). *Web analytics: An hour a day*. Indianapolis, IN: Sybex.

Kennedy, A. (2014). Introduction for teachers. *SHOW.ME*. Retrieved from: http://www.show.me.uk/editorial/1155-introduction-for-teachers Accessed June 10, 2017.

Kershaw, S. (2013, November 9). CSS Nav Bar Padding and Overflow [Msg 1]. Message posted to https://stackoverflow.com/questions/19871166/css-nav-bar-padding-and-overflow

Khoon, L. C., & Ramaiah, C. K. (2008). An overview of online exhibitions. *DESIDOC Journal of Library & Information Technology*, *28*(4), 7.

Knell, S. J. (2003). The shape of things to come: Museums in the technological landscape. *Museum and Society*, 1(3), 132-146. Retrieved from: http://hdl.handle.net/2381/53

Krippendorff, K. (1989). Content analysis. In E. Barnouw, G. Gerbner, W. Schramm, T. L. Worth, & L. Gross (Eds.), *International encyclopedia of communication*, 1, 403-407. New York, NY: Oxford University Press. Retrieved from http://repository.upenn.edu/asc_papers/226n

Krippendorff, K. (2004). Reliability in Content Analysis: Some Common Misconceptions and Recommendations. *Human Communication Research*, 30 (3), 411-433. http://dx.doi.org/10.1111/j.1468-2958.2004.tb00738.x

Krug, S. (2014). *Don't make me think revisited: A common sense approach to web and mobile usability*. Berkeley, CA: New Riders.

Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and best practices in content analysis. *Journal & Mass Communication Quarterly*, 92(4), 791-811.

Leedy, Paul D., & Ormrod, Jeanne E. (Eds.). (2010). *Practical research: Planning and design*. Saddle River, NJ: Pearson.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... & Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22-31. doi: 10.1145/1629607.1629613

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... & Wolff, S. S. (1997). The past and future history of the Internet. *Communications of the ACM*, 40(2), 102-108. doi: 10.1145/253671.253741

Lovelace, A. C. (1842). Translator's notes to an article on Babbage's Analytical Engine. *Scientific Memoirs*, 3, 691-731.

Madrigal, A. C. (2011). IBM's First 100 Years: A Heavily Illustrated Timeline. *The Atlantic*. Retrieved from: https://www.theatlantic.com/technology/archive/2011/06/ibms-first-100-years-a-heavily-illustrated-timeline/240502/

Malraux, A. (1967). *Museum without walls*. (S. Gilbert & F. Price, Trans.). Garden City, NY: Doubleday & Company, Inc. (Original work published 1965)

Mannoni, B. (1996). Bringing museums online. *Communications of the AMC*, 39(6), 100-105. doi: 10.1145/228503.228524

Mannoni, B. (1997). A virtual museum. *Communications of the AMC*, 40(9), 61-62.

Maranto, J. V. (2015, February, 2). *History of the museums* [Video file]. Retrieved from http://ed.ted.com/lessons/why-do-we-have-museums-j-v-maranto#watch

Marty, P. F. (2008). Museum websites and museum visitors: Digital museum resources and their use. *Museum Management and Curatorship*, 23(1), 81-99. http://dx.doi.org/10.1080/09647770701865410

Maxwell, J. A. (2009). Designing a qualitative study. In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (214-253). Thousand Oaks, CA: SAGE Publication, Inc.

Mayring, P. (2000). Qualitative content analysis. *Forum qualitative sozialforschung/forum: Qualitative social research*, 1(2). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385

McDonald, M. (2005). The museum and the web: Three case studies. Master's thesis, American Studies, University of Virginia, Charlottesville, Virginia.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Morville, P. (2004). A brief history of information architecture. Information architecture: Designing information environments for purpose. Retrieved from: http://semanticstudios.com/pdfs/historia.pdf Accessed July, 20, 2017.

Morville, P. (2011, August). *Information architect*. Semantic Studios. Retrieved from http://semanticstudios.com/information_architect/ Accessed July, 20, 2017.

Morville, P. & Rosenfeld, L. (2006). *Information architecture for the world wide web* (3rd ed.). Sebastopol, CA: O'Reilly Media Inc. Retrieved from: http://proquest.safaribooksonline.com.login.ezproxy.library.ualberta.ca/book/web-development/0596527349/firstchapter.

Müller, K. (2002). Museums and virtuality. *Curator: The Museum Journal*, 45(1), 21-33. Doi: 10.1111/j.2151-6952.2002.tb00047.x

Nelson, T. H. (1987). *Literary machines 93.1.: the report on, and of, project Xanadu concerning word processing, electronic publishing, hypertext, thinker toys, tomorrow's intellectual* *revolution, and certain other topics including knowledge, education and freedom*. Sausalito, CA: Mindful Press. (Originally published in 1980).

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. *CHI 92' Conference*, 373-380. doi: 10.1145/142750.142834

Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. *In CHI '94 Human Factors in Computing Science*, 152-158. doi: 10.1145/191666.191729

Nielsen, J. (1994b). *Usability Engineering*. Cambridge, MA: Academic Press Inc.

Nielsen, J. (1994c). Usability inspection methods. *In CHI '94 Conference Companion on Human Factors in Computing Systems*, 413-414. doi: 10.1145/259963.260531

Nielsen, J. (1995a, January 1). *How to conduct a heuristic evaluation*. Retrieved from: https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/

Nielsen, J. (1995b, January 1). *Severity ratings for usability problems*. from: https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

Nielsen, J. (2012, January 4). *Usability 101: Introduction to usability*. Retrieved from https://www.nngroup.com/articles/usability-101-introduction-to-usability/ Accessed July 26, 2017.

Nielsen, J. K. (2015). The relevant museum: Defining relevance in museological practices. *Museum Management and Curatorship*. 30(5), 364-378. DOI: 10.1080/09647775.2015.1043330 Accessed June 26, 2017.

Nielsen, J., & Mack, R. L. (1994). *Usability inspection methods*. New York: Wiley.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *CHI '90 Proceedings, April*. 249-256. doi: 10.1145/97243.97281

Novick, D. G. & Hollingsed, T. (2007). Usability inspection methods after 15 years of research and practice. *Departmental Papers (CS).* Paper 16. http://digitalcommons.utep.edu/cs_papers/16

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. Stanford InfoLab. Retrieved from: http://ilpubs.stanford.edu:8090/422/

Pallas, J. & Economides, A. A. (2008). Evaluation of art museums' web sites worldwide. Information *Services and Use*, 28(1), 45-57. doi: 10.3233/ISU-2008-0554

Paolini, P., Barbieri, T., Loiudice, P., Alonzo, F., Zanti, M., & Gaia, G. (2000). Visiting a museum together: How to share a visit to a virtual world. *Journal of the Association for Information Science and Technology*, 51(1), 33-38. doi: 10.1002/(SICI)1097-4571(2000)51:1<33::AID-ASI6>3.0.CO;2-I

Parandjuk, J. C. (2010). Using information architecture to evaluate digital libraries. *The Reference Librarian*, 51(2), 124-134. http://dx.doi.org/10.1080/02763870903579737

Piacente, M. (1996). Surf's up: Museums and the world wide web. MA Research Paper, Museum Studies Program. University of Toronto.

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of applied communication research*, 27(1999), 258-284. http://dx.doi.org/10.1080/00909889909365539

Pulh, M., & Mencarelli, R. (2015). Web 2.0: Is the museum-visitor relationship being redefined?. *International Journal of Arts Management*, 18(1), 43-51. Retrieved from: https://search.proquest.com/docview/1717299064?accountid=33837

Resmini, A., & Rosati, L. (2012). A brief history of information architecture. *Journal of Information Architecture*, 3(2). Retrieved from: http://journalofia.org/volume3/issue2/03-resmini/jofia-0302-03-resmini.pdf

Riley-Huff, D. A. (2012). Web accessibility and universal design: A primer on standards and best practices for libraries. *Making Libraries Accessible: Adaptive Design and Assistive Technology*, 48(7), 29-35.

Roller, M. R., & Lavrakas, P. J. (2015). *Applied qualitative research design: A total quality framework approach*. New York: The Guilford Press.

Rosenfeld, L., Morville, P. & Arango, J. (2015). *Information architecture: For the web and beyond* (4th ed.). Sebastopol, CA: O'Reilly Media.

Sauro, J. (2010, August 31). *6 things you didn't know about heuristic evaluations*. Retrieved from: https://measuringu.com/he/

Sauro, J. (2012a, August 14). *The value of multiple evaluators in heuristic evaluations*. Retrieved from: https://measuringu.com/he-multiple/

Sauro, J. (2012b, September 6). *How effective are heuristic evaluations?* Retrieved from https://measuringu.com/effective-he/

Sauro, J. (2016, February 2). *Managing false positives in ux research*. Retrieved from: https://measuringu.com/false-positives/

Schreier, M. (2012). *Qualitative content analysis in practice*. London, UK: Sage Publications Ltd.

Schreier, M. (2014). Qualitative content analysis. In Flick, U. (Ed.), *The SAGE handbook of qualitative data analysis* (170-183). London, UK: SAGE Publications Ltd. doi: 10.4135/9781446282243.n12

Schweibenz, W. (1998). The "virtual museum": New perspectives for museums to present objects and information using the internet as a knowledge base and communication system. *Internationales Symposium für Informations – wissenschaft ISI*. 185-200.

Schweibenz, W. (2004). The development of virtual museums. *ICOM News*, No. 3. Retrieved from: http://cool.conservation-us.org/icom/fileadmin/user_upload/pdf/ICOM_News/2004-3/ENG/p3_2004-3.pdf

Sears, A. (1997). Heuristic walkthroughs: Finding the problems without the noise. *International Journal of Human-Computer Interaction*, 9(3), 213-234. http://dx.doi.org/10.1207/s15327590ijhc0903_2

Sellmer, M. (2012). *The Information Architecture of Virtual Museums*. Unpublished manuscript, Library of Information Studies, University of Alberta, Edmonton, Alberta.

Simon, S. J. (2008). Information architecture for digital libraries. *First Monday*, 13(12). http://dx.doi.org/10.5210/fm.v13i12.2183

The Smithsonian Institute (2017). Our history Retrieved from https://www.si.edu/about/history Accessed July 10, 2017.

Smithsonian Learning Lab. (2017). About the Smithsonian Learning Lab. Retrieved from: https://learninglab.si.edu/about Accessed June 10, 2017.

Spencer, D. (2011). *A Practical Guide to Information Architecture*. Five Simple Steps.

Srinivasan, R., Boast, R., Furner, J., & Becvar, K. M. (2009). Digital museums and diverse cultural knowledges: Moving past the traditional catalog. *The Information Society*, 25(4), 265-278. http://dx.doi.org/10.1080/01972240903028714

Stuedahl, D. (2007). Convergence, museums and digital cultural heritage. *The Ambivalence of Convergence*, edited by Tanja Storsul and Dagny Stuedahl, NORDICOM Gothenborg, 129-143. Retrieved from: https://www.researchgate.net/profile/Dagny_Stuedahl/publication/261830498_Convergence_museums_and_digital_cultural_heritage/links/02e7e5358f25ac7b4d000000.pdf

Styliani, S., Liarokapis F., Kotsakis K., & Patias P. (2009). Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage*. 10: 520-528. https://doi.org/10.1016/j.culher.2009.03.003

Sviličić, N. (2010). Creation and typology definition of the museum on the Internet. *Collegium Antropologicum*, *34*(2), 587-594. Retrieved from: http://hrcak.srce.hr/56487

Tam, L. W. H., & Robertson, A. C. (2002). Managing change: Libraries and information services in the digital age. *Library Management*, 23(8/9), 369–377. https://doi.org/10.1108/01435120210439843

Tashakkori, A. & Teddlie, C. (2009). In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (283-317). Thousand Oaks, CA: SAGE Publication, Inc.

Teather, L. (1998). A museum is a museum is a museum… Or is it?: Exploring museology and the web. *Museums and the Web 1998*. Retrieved from: http://www.museumsandtheweb.com/mw98/papers/teather/teather_paper.html

The Open Graph Protocol (2017). Introduction. Retrieved from: http://ogp.me Accessed June 10, 2017.

Toub, S. (2000). Evaluating information architecture: A practical guide to assessing web site organization. *Argus Center for Information Architecture*.

Tracy, S. J. (2013). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact.* Chichester, UK: Wiley-Blackwell.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. Retrieved from http://www.jstor.org/stable/2251299

Usability.gov (2013). Information architecture basics. Retrieved from https://www.usability.gov/what-and-why/information-architecture.html# Accessed July 20, 2017.

English Oxford Dictionary (n.d.). User [Def. 1]. Retrieved from: https://en.oxforddictionaries.com/definition/user

Vicars-Harris, O. (2016, November). *Digital Strategy*. Retrieved from: http://www.npg.org.uk/assets/files/pdf/strategic-plan/NPG_Digital_Strategy_Digest_v4_1.pdf Accessed June 10.

Encyclopædia Britannica Online. (2017). Virtual museum. Retrieved from https://www.britannica.com/topic/virtual-museum Accessed June 30, 2017.

W3schools.com (n.d). *CSS Overflow*. Retrieved from: https://www.w3schools.com/css/css_overflow.asp Accessed June 24, 2017.

Walsh, D., Clough, P., & Foster, J. (2016). User Categories for Digital Cultural Heritage. In *Proceedings of 1st International Workshop on Accessing Cultural Heritage at Scale.* June

22, 2016, Newark NJ. Retrieved from https://repository.edgehill.ac.uk/7790/1/paper6.pdf Accessed July 18, 2017.

Wang, R. G., Dong, H. L., Wu, M. W., & Wu, Q. F. (2014). Research and design of digital museum based on virtual reality. *Advanced Materials Research*, Vol. 926, 2516-2520. doi: 10.4028/www.scientific.net/AMR.926-930.2516

Web Gallery of Art (2017a). Entrance Page. Retrieved from: http://www.wga.hu Accessed June 10, 2017.

Web Gallery of Art (2017b). Homepage. Retrieved from: http://www.wga.hu/index1.html Accessed June 10, 2017.

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994, June). The cognitive walkthrough method: A practitioner's guide. In *Usability inspection methods* (pp. 105-140). John Wiley & Sons, Inc..

WikiArt. (2017). About. Retrieved from: https://www.wikiart.org/en/about Accessed June 10, 2017.

Wilkes, M. V. (2002). Charles Babbage and his world. *Notes and Records*, 56(3), 353-365. doi: 10.1098/rsnr.2002.0188

Williams, J. B. (2017). Going Online: The Internet. In *The Electronics Revolution* (pp. 205-214). Springer International Publishing. https://doi.org/10.1007/978-3-319-49088-5_22

Wilson, V. (2011). Research methods: content analysis. *Evidence based library and information practice*, 6(4), 177-179.

Wodtke, C. & Govella, A. (2009). *Information Architecture: Blueprints for the Web* (2nd ed.). A. Govella (ed.). Thousand Oaks, CA: New Riders Publishing.

Woolrych, A., Hornbæk, K., Frøkjær, E., & Cockton, G.. (2011). Ingredients and meals rather than recipes: A proposal for research that does not treat usability evaluation methods as indivisible wholes. *International Journal of Human-Computer Interaction*, 27(10), 940-970. http://dx.doi.org/10.1080/10447318.2011.555314

Wurman, R. S. (1997). *Information Architects*. New York City, NY: Watson-Guptill.

Wyman, B., Smith S., Meyers D., & Godfrey, M. (2011). Digital storytelling in museums: Observations and best practices. *Curator,* 54(4), 461–468. doi: 10.1111/j.2151-6952.2011.00110.x

Zhou, M., Geng, G., & Wu, Z. (2012). *Digital preservation technology for cultural heritage*. Springer Science & Business Media. doi: 10.1007/978-3-642-28099-3

Appendices

Appendix A – Glossary

Digital Museums: Are online environments that use different technologies (like 3D graphics and multimedia) to present collections of objects with contextual information in order to create an experience for users (Foo, 2008; Schweibenz, 1998; Schweibenz, 2004; Styliani et al., 2009; Zhou et al., 2012).

Criteria: “a characterizing mark or trait” (Merriam-Webster Dictionary, n.d., Criterion).

Information Architecture: The definition for this study is from *Information architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville, and Jorge Arango (2015). It is:

- The structural design of shared information environments
- The synthesis of organization, labeling, search, and navigation systems within digital, physical, and cross-channel ecosystems
- The art and science of shaping information products and experiences to support usability, findability, and understanding
- An emerging discipline and community of practice focused on bringing principles of design and architecture to the digital landscape

(Rosenfeld et al., 2015, 24)

Labeling System: “Labeling is a form of representation… we use labels to represent larger chunks of information in our information environments.” (Rosenfeld et al., 2015, 133)

Navigation System: Navigation systems can include hypertext links, global navigation menus, and search engines that allow users to move from page to page within websites (ABC-CLIO, 2012).

Organization Systems: “The main ways of categorizing or grouping a site’s content (e.g., by topic, by task, by audiences, or by chronology). Also known as taxonomies and hierarchies. Organization systems are composed of organization schemes and organization structures. An organization scheme defines the shared characteristics of content items and influences theological grouping of those items. An organization structure defines the types of relationships between content items and groups.” (Morville & Rosenfeld, 2006, p. 58)

Principle: This is “A natural law forming the basis for the construction or working of a machine” (Oxford English Dictionary, n.d.). The “principles” for this study are the definitions and descriptions of information architecture components used for the organization, navigation, searching, labeling, and vocabulary in website design. They can be associated with either the structure and/or language of a website, but they are not philosophies, they are established components used in real world websites that enable the organization of and access to information.

Search System: Software systems that help users find information on websites by selecting predetermined search terms or by entering keywords and clicking on the search button (ABC-CLIO, 2012).

User: A human who uses or interacts with something (Oxford English Dictionary, n.d.,”User”).

Vocabulary Systems: “These systems allow you to structure and map languages so that people can more easily find information.” (Rosenfeld et al., 2015, 309)

Appendix B – Initial Coding Frame

The unit of analysis for this study is four general knowledge information architecture books. They are:

- *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella (coded CW)
- *A Practical Guide to Information Architecture* by Donna Spence (coded DS)
- *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango (coded RMA)
- *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert (coded AC)

The study will analyze each chapter in all the books, however, the preface, table of contents, coda, appendices, index(s), footnotes and endnotes, and reference/bibliography sections will not be included.

Information in tables and images will be excluded from analysis.

Direct quotes from secondary sources will not be coded.

This study will use a formal unit of coding. All coding will consist of at least two words up to a maximum of four sentences.

Only information and text that is directly pertinent to the research question (below) will be coded.

Research Question:

“What are the information architecture principles present in the select information architecture literature?”

The focus of this content analysis study is the identification of **information architecture principles**.

Information architecture is the structure and language that make websites understandable for the user and lets them find the information they are looking for (The Information Architecture Institutes, 2017). This is made up of individual structure or language components organized into the categories:

- Organization systems
- Navigation systems
- Search systems
- Labeling systems
- Vocabulary

The **principles** for this study are the definitions and descriptions of information architecture methods used for the organization, navigation, searching, labeling, and vocabulary in website design. They can be associated with either the structure and/or language of a website, but they are not philosophies, they are established components used in real world websites that enable the organization of and access to information.

Materials that will not be coded include anything outside the scope of the research question. For example, information about graphic design, website users, wireframes, etc. will not coded because they do not directly relate to the "information architecture principles."

Not every time the words organization, navigation, search, label, or vocabulary appear will the information be coded, it must be directly associated with a principle.

Information about the different types of user or how users use a website will not necessarily be coded, only if the text discuss the information architecture principles that they used.

Information about website context and how, for example, business goals influence site design will not be coded unless a principle is mentioned. This information will be discussed in further detail in the literature review.

Additionally, discussions about website content (image and text presentation, etc.) will not be coded unless a principle is present.

To provide context for the reader the page number and book code will be provided, so that they may read the excerpted quote.

Coding Instructions:

When coding the text in the pilot study, all categories will be coded with the first letter of the main category it belongs to (e.g., O, N, S, etc.).

When creating subcategories they will be coded with the first letter of the main category and then the first letter of the subcategories. For example under the search system one sub category is advanced search so its code would be SA.

If there are subcategories within a subcategory these will be labeled with the first letter of all the categories it belongs to. For example, an indexing thesaurus is a subcategory of thesaurus, which is a subcategory of vocabulary, so its code would be VTI.

The main categories of analysis (subcategories will be subsumed during a pilot study) are:

Category	**Definition**	**Example**	**Code**
Organization System	"Organization systems are composed of organization schemes and organization structures. An organization scheme defines the shared characteristics of content items and influences theological grouping of those items. An organization structure defines the	"Exact or 'objective' organization schemes divide information into well-defined and mutually exclusive sections. For example, country names are usually listed in alphabetical order." (Rosenfeld et al., 2015, 105)	O

	types of relationships between content items and groups." (Morville & Rosenfeld, 2006, p. 58)		
Navigation System	"The use of hypertext links, icons, menu options, and search engines displayed on a Web page to move to other resources available on the Internet or to other pages within the same Web site."(ABC-CLIO, Navigation, 2012)	*"Local navigation systems*: Primary navigation systems that help users understand where they are and where they can go within an information environment." (Rosenfeld et al., 2015, 91)	N
Search System	"How search systems are powered by classification schemes, and how users use them to obtain information resources relevant to an information need/search." (The Information Architecture Institute, 2017)	"There are the guts of the search engine itself; aside from tools for indexing and spidering, there are algorithms for processing your query into something the software can understand, and for ranking the results." (Rosenfeld et al., 2015, 217)	S
Labelling System	"Labeling is a form of representation… we use labels to represent larger chunks of information in our information environments." (Rosenfeld et al., 2015, 133)	"Well labels are often the most obvious way to clearly show the user your organization and navigation schemes across multiple systems and contexts." (Rosenfeld et al., 2015, 134)	L
Vocabulary Systems	"These systems allow you to structure and map languages so that people can more easily find information." (Rosenfeld et al., 2015, 309)	"At its simplest, a controlled vocabulary is a list of equivalent terms in the form of a synonym ring, or a list of preferred terms in the form of an authority file." (Rosenfeld et al., 2015, 271)	V
Miscellaneous	This field is included so that all coded information can be categorized. Any information that does not belong in one of the five other categories will be found here.		M

Coding Rules:

If there is overlap between subcategories within the same category:

- First see if the overlap can be separated in to two units of coding
- If it cannot be separated, the author must interpret which subcategory is the main topic of that unit of text.

All categories will be mutually exclusive.

Appendix C – Content Analysis Report Form

Literature excerpt	Brief Summary	Code	Category	Subcategory	Researcher Memos
"The foundation of many good information architectures is a well designed hierarchy... The mutually exclusive subdivisions and parent-child relationships of hierarchies are simple and familiar." (Rosenfeld et al., 2015, 117)	Organization structure - hierarchy design	OH (this stands for Organization System "O", and hierarchy "H")	Organization System	Hierarchy	This is an example of how this form will be filled out. Coding is subject to changes during the initial analysis and pilot study.

Appendix D – Heuristic Evaluation Report Form

Unit ID	Date	Museum	Museum Classification	Location in website	Problem or Bug	Heuristic	Researcher Notes	True/False
1	09/04/2017	The Louvre	Brochure	Search results (from collection page)	Searched for "Venus de Milo," and it was not the first results displayed (it was the fourth result on the list)	Search precision	This is an example of how forms will be filled out as the heuristic evaluations take place. This museum will not be studied as it was part of the pilot study.	

Appendix E – Museum Sample List

Brochure Museums:

- Art Gallery of Ontario
    - https://www.ago.net/
- National Portrait Gallery
    - http://www.npg.org.uk
- Cleveland Museum of Art
    - http://www.clevelandart.org

Content Museums:

- Art UK
    - https://artuk.org
- Discover Islamic Art
    - http://www.discoverislamicart.org/
- WikiArt
    - https://www.wikiart.org

Learning Museums:

- Smithsonian Learning lab
    - https://learninglab.si.edu
- SHOW.ME
    - http://www.show.me.uk
- Web Gallery of Art
    - http://www.wga.hu/index.html

Appendix F – Categories and Subcategories

Categories and subcategories of content analysis (created through subsumption):

Organization System:

- Organization scheme
    - Exact organization scheme
        - Alphabetical schemes
        - Chronological schemes
        - Geographical schemes
    - Ambiguous organization scheme
        - Topical organization scheme
        - Task oriented organization scheme
        - Audience specific organization scheme
        - Metaphor-driven organization scheme
    - Hybrid organization scheme
    - Miscellaneous
- Organization structure
    - Hierarchy structure
    - Polyhierarchical structure
    - Hypertext structure
    - Database oriented
    - Social tagging
    - Hybrid structure
- Miscellaneous

Labeling System:

- Types of Labels
    - Textual labels
        - Contextual links
        - Headings

        - Navigation labels
        - Index terms
    - Icon labels
- Labeling consistency
- Miscellaneous

Navigation Systems:

- Global navigation
- Local navigation
- Contextual navigation
- Breadcrumb navigation
- Hypertext navigation
- Navigation tools
    - Sitemaps
    - Indexes
    - Guides
    - Configurators/wizards
- Advanced navigation
    - Personalization
    - Customization
    - Visualization
    - Social Navigation
- Miscellaneous

Search Systems

- Search algorithms
    - Types of algorithms
    - Recall and precision
- Search interface
    - Search interface components
    - Advanced search

- Search results
    - Displaying results
        - Format
        - Information displayed for retrieved items
        - Miscellaneous
    - Sorting results
        - Alphabetically
        - Chronologically
    - Ranking results
        - Popularity
        - Users' or experts' ratings
        - Pay-for-placement
        - Miscellaneous
    - "Best bets"
    - Hybrid search results
    - Additional actions
        - Save search
        - Select subset of results
        - Narrowing results down
        - Repeating/new search
    - Miscellaneous
- Query builders
    - Spell checkers
    - Phonetic tools
    - Stemming tools
    - Natural language processing tools
    - Autocomplete/Autosuggestions
- Query languages
- Indexing content for searching
    - Search zones
        - Navigation and Destination webpages

        - Indexing by topic
        - Indexing for a specific audience
        - Indexing recent content
        - Indexing full-text
    - Miscellaneous
- Miscellaneous

Vocabulary Systems:

- Metadata
    - Structural metadata
    - Descriptive metadata
    - Administrative metadata
- Controlled Vocabulary
    - Synonym rings
    - Authority files
    - Classification schemes
- Thesauri
    - Classic thesaurus
    - Indexing Thesaurus
    - Searching Thesaurus
    - Thesaurus standards
    - Thesaurus terms
    - Miscellaneous
- Semantic Relationships
    - Hierarchical relationships
    - Polyhierarchical relationships
    - Equivalence relationships
    - Associative relationships
- Faceted Classification
- Miscellaneous

Appendix G – Final Coding Frame

Coding Frame

The units of analysis for this study are four general knowledge information architecture books. They are:

- *Information Architecture: Blueprints for the Web* by Christina Wodtke and Austin Govella (coded 1)
- *A Practical Guide to Information Architecture* by Donna Spencer (coded 2)
- *Information Architecture for the Web and Beyond* by Louis Rosenfeld, Peter Morville and Jorge Arango (coded 3)
- *How to Make Sense of Any Mess: Information Architecture for Everybody* by Abby Covert (coded 4)

The study will analyze each chapter in all books, however, the preface, table of contents, coda, appendices, index(s), footnotes and endnotes, and reference/bibliography sections will not be included.

This study will use a formal unit of coding. All coding will consist of at least one word to a maximum of four sentences.

To provide context for the readers, the book codes (in its own column), and page numbers will be provided, so that they may read the expanded quote and surrounding text.

Multiple instances of each principle will be coded. This is so that the frequency of the principles can be established, so that quantitative methods can be used to support (or provide evidence against) qualitative findings.

Relevant and Irrelevant Material

Only information and text that is directly pertinent to the research question (below) will be coded.

Research Question:

"What are the information architecture principles present in the select information architecture literature?"

The focus of this content analysis study is the identification of **information architecture principles**.

Information architecture is the structure and language that make websites understandable for the user and lets them find the information they are looking for (The Information Architecture Institutes, 2017). This is made up of individual structure or language components organized into the categories:

- Organization systems
- Navigation systems
- Search systems
- Labeling systems
- Vocabulary

The **principles** for this study are the definitions and descriptions of information architecture components used for the organization, navigation, searching, labeling, and vocabulary in website design. They can be associated with either the structure and/or language of a website, but they are not philosophies, they are established components used in real world websites that enable the organization of and access to information.

Materials that will not be coded include anything outside the scope of the research question. For example, information about graphic design, website users, wireframes, etc. will not coded because they do not directly relate to the "information architecture principles." This includes the discussion of principles not related to IA (for example real world architecture or graphic design).

Not every time the words organization, navigation, search, label, or vocabulary appear will the information be coded, it must be directly associated with a principle.

Information about the different types of user or how users use a website will not necessarily be coded, only if the text discuss the information architecture principles that they used. This applies to user action and not website structure and/or labeling (the principles).

Information about website context and how, for example, business goals influence site design will not be coded unless a principle is mentioned. This information will be discussed in further detail in the literature review.

Information in tables and images as well as the associated descriptions will be excluded from analysis.

Quotes from secondary sources will not be coded.

Additionally, discussions about website content (image and text presentation, etc.) will not be coded unless a principle is present.

Coding Rules and Instructions

The categories and subcategories are coded numerically, from 1 (Organization System) to 115 (a Miscellaneous category under Vocabulary Systems). The content analysis form will provide a unique identifier for each entry.

Coding will be done numerically not texturally – to facilitate coding data analysis in Microsoft Excel. This includes the categories, subcategories, source material, and unit IDs.

If there is overlap between subcategories:

- First see if the overlap can be separated in to two units of coding.

- If it cannot be separated, the author must interpret which subcategory is the main topic of that unit of text. The main topic is generally the more specific principle discussed.
- If there are two principles in the same sentence, especially if they are associated with the same category, check to see if there is a "Hybrid" subcategory (it would be belong there).

Examples:

- Grappling with these local navigation issues can make creating global navigation systems look easy. (Rosenfeld et al., 2015, 188)
    - This sentence can (and must to retain mutual exclusiveness) be separated into two units of coding. One about local navigation and the other about global navigation.

- "Or, you might ignore synonym rings for initial searches but provide the option to 'expand your search to include related terms' if there were few or no results." (Rosenfeld et al., 2015, 275)
    - You can code this under Synonym rings and Repeating/new search.

- "Organization systems present the site's information to us in a variety of ways, such as content categories that pertain to the entire campus (e.g., the top bar and its "Academics" and "Admission" choices), or to specific audiences (the block on the middle left, with such choices as "Future Students" and "Staff")." (Rosenfeld et al., 2015, 82)
    - This is an example of a "Hybrid" category. The website being discussed has both topical ("Admission") and audience ("Staff") organization schemes.

Warning!

Information architecture principles are closely connected, to the point that the same terms are used in different categories. The coder must be aware of these duplicate instances and use judgment and surrounding textual context to determine what category it belongs in or code the text excerpt in both categories (as long as the categories remain mutually exclusive). Additionally, some terms discuss an information architecture principle, but those terms can also be used to discuss something that is not a principle.

Examples:

- Contextual links is in both labeling and navigation systems.
- The term index is used in navigation, searching, and vocabulary.
- Hierarchy can refer to a semantic relationship and organization structure.
- There are different "Hybrid" categories.
- The appearance of synonym can be classified under "Synonym ring" or "Equivalence relationships" (associated with synonym management).
- Classification scheme can mean either relationships between preferred terms (as defined in this study) or some information architects use this term instead of organization scheme.
- This is true for the meaning of taxonomies. In the pilot study this term was encountered when discussing both the organization hierarchy structure and classification schemes.
- The term sitemaps can refer to the website supplemental navigation principle or a research design tool.

If a chunk of text can be coded into multiple categories, it needs to be coded separately in the coding form (for data analysis purposes).

All categories must remain mutually exclusive.

The categories and subcategories definitions:

Main Category	Subcategories	Definition	Example	Code
Organization System		"Organization systems are composed of organization schemes and organization structures. An organization scheme defines the shared characteristics of content items and influences theological grouping of those items. An organization structure defines the types of relationships between content items and groups." (Morville & Rosenfeld, 2006, p. 58)	"Organization systems present the site's information to us in a variety of ways, such as content categories that pertain to the entire campus (e.g., the top bar and its "Academics" and "Admission" choices), or to specific audiences (the block on the middle left, with such choices as "Future Students" and "Staff")." (Rosenfeld et al., 2015, 82)	1
	Organization scheme	"An organization scheme defines the shared characteristics of content items and influences the logical grouping of those items." (Rosenfeld et al., 2015, 103)	(See subcategories)	2
	Exact organization scheme	"Exact or 'objective' organization schemes divide information into well-	(See subcategories)	3

				defined and mutually exclusive sections.” (Rosenfeld et al., 2015, 105)		
			Alphabetical scheme	“An alphabetical organization scheme is the primary organization scheme for encyclopedias and dictionaries.” (Rosenfeld et al., 2015, 105)	“For example, country names are usually listed in alphabetical order. If you know the name of the country you are looking for, navigating the scheme is easy. “Chile” is in the Cs, which are after the Bs but before the Ds.” (Rosenfeld et al., 2015, 105)	4
			Chronological scheme	The organization of information by date (any date – date published – historical event date, etc.).	“History books, magazine archives, diaries, and television guides tend to be organized chronologically.” (Rosenfeld et al., 2015, 106)	5
			Geographical scheme	The organization of information by place or location.	Choosing your local weather by entering your postal code is an example of organizing information by geography.	6
		Ambiguous organization scheme		“Ambiguous or “subjective” organization schemes divide information into categories that defy exact definition.” (Rosenfeld et al., 2015, 108)	(See subcategories)	7
			Topical scheme	“Organizing information by subject or topic is one of the most useful and challenging approaches. Newspapers are organized topically, so if you want to see the scores from yesterday’s game, you know to turn to the sports section.” (Rosenfeld et al., 2015, 109)	For example, the Brick website organizes information by topic. If you are looking for a sofa you look under the topic “Furniture” and select “Sofas” from the dropdown menu (it’s beneath “Living Room”).	8
			Task oriented scheme	“Task-oriented schemes organize content and applications into collections of processes, functions, or tasks. These schemes are appropriate when it’s possible to anticipate a limited number of high-priority tasks that users will want to perform.” (Rosenfeld et al., 2015, 110)	If the navigation headings are verbs, then in general, the website is using some form of task oriented organization scheme. Like “Explore, Visit, Learn, etc.” The Princeton public library uses the action words ”Find It, Attend, Connect, Explore, and Get to know us.”	9
			Audience-specific scheme	“Audience-oriented schemes break a site into smaller, audience-specific mini-sites, thereby allowing for clutter-free pages that present only the options of interest to that particular audience.” (Rosenfeld et al., 2015, 112)	The Metropolitan Museum of Art website organizes information that people want to “Learn” (the subject heading this is found under) about by audience. This includes “Kids & Families,” “Teens,” “Educators, “ etc.	10
			Metaphor-driven scheme	“Metaphors are commonly used to help users understand the new by relating it to the familiar. You need not look further than your desktop computer with its folders, files, and trash can or recycle bin for an example.” (Rosenfeld et al., 2015, 113)	One of the most popular examples would be the Apple iBooks app, which displays your books arranged on a virtual shelf.	11
		Hybrid organization scheme		“…hybrid scheme includes elements of audience-specific, topical, metaphor-based, task-oriented, and alphabetical organization schemes.” (Rosenfeld et al., 2015, 114)	This is a very common option. Lets look at the Metropolitan Museum of Art website again, we know that they organize information by audience, but they also use the chronologically scheme, “Exhibitions” (with current, upcoming and past options).	12
		Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here.		13
	Organization structure			“An organization structure defines the types of relationships between content items and groups.” (Rosenfeld et al., 2015, 105)	(See subcategories)	14

		Hierarchy structure		"The mutually exclusive subdivisions and parent–child relationships of hierarchies are simple and familiar. We have organized information into hierarchies since the beginning of time. Family trees are hierarchical." (Rosenfeld et al., 2015, 117)	Most websites have a hierarchical organization structure. Their global navigation headings are the "parents"; with the content organized underneath the "children" There can be shallow hierarchies (only one or two levels below the "parent" heading) or deep hierarchies (multiple layers).	15
		Polyhierarchical structure		"Within a single organization scheme, you will need to balance the tension between exclusivity and inclusivity. Hierarchies that allow cross-listing are known as polyhierarchical". (Rosenfeld et al., 2015, 118)	In the Edmonton Public Library access to "eBooks" is available under the navigation headings "Browse" and "Digital Content." This is a cross listing of information.	16
		Hypertext structure		"Hypertext is a highly nonlinear way of structuring information. A hypertext system involves two primary types of components: the items or chunks of information that will be linked, and the links between those chunks." (Rosenfeld et al., 2015, 126)	Purely hypertext structures are uncommon (often combined with other organization structures – see hybrid structure below). The website New 7 Wonders of the World uses hypertext to connect all the pages and access different section of the same page (moving the user down the page depending on the links clicked).	17
		Database structure		"In relational database structures, data is stored within a set of relations or tables. Rows in the tables represent records, and columns represent fields. Data in different tables may be linked through a series of keys." (Rosenfeld et al., 2015, 122-123)	Websites with very large collection of information may use a database structure (which is better suited for storing large quantities of data) like the online digital library Artstor.	18
		Social tagging		"Users tag objects with one or more keywords. These tags can be informally supported in text fields, or they can be provided for with bespoke fields in the formal structure of content objects." (Rosenfeld et al., 2015, 127)	"LinkedIn allows users to "endorse" their professional contacts as possessing certain individual professional skills. These endorsements are in effect tags: they allow users to describe their business contacts in a granular way that informs how the system groups them with similar people. (Rosenfeld et al., 2015, 128)	19
		Hybrid structure		This is the combination of two or more organization structures.	Many websites use hybrid organization structures. They have a hierarchical structure but also use social tagging and/or hypertext to connect "children" pages.	20
	Miscellaneous			This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here.		21
Labeling System				"Labeling is a form of representation… we use labels to represent larger chunks of information in our information environments." (Rosenfeld et al., 2015, 133)	"Well labels are often the most obvious way to clearly show the user your organization and navigation schemes across multiple systems and contexts." (Rosenfeld et al., 2015, 134)	22
	Types of Labels			"Labels should educate people about new concepts and help them quickly identify familiar ones." (Rosenfeld et al., 2015, 135)	(See Subcategories)	23
		Textual Labels		Textual labels are the words or a word that represent sections of information within a website (these can either be static or links). It is important to keep in mind the users of the website when developing these terms.	(See subcategories)	24
			Contextual links	"Labels describe the hypertext links within the body of a document or chunk of information, and naturally occur within the descriptive context of their surrounding text. Contextual	If you have ever encountered a link within an article (like in Wikipedia) that takes you to a similar or related webpage then you have encountered a contextual link.	25

				links are easy to create and are the basis for the exciting interconnectedness that drives much of the Web's success." (Rosenfeld et al., 2015, 141)		
			Headings	"Headings, as shown in, are often used to establish a hierarchy within content. Just as in a book, where headings help us distinguish chapters from sections, they also help us determine a site's subsites, or differentiate categories from subcategories." (Rosenfeld et al., 2015, 144)	The Louvre website uses headings to define different subsection of the site – "Plan your Visit," "Activities & Tours," Exhibitions & Events," etc. Each heading has multiple subheadings organized under it (this also represents the hierarchy of the site).	26
			Navigation labels	"Users rely on a navigation system to behave "rationally" through a consistent location and look; [navigation] labels should be no different. Effectively applied labels are integral to building a sense of familiarity, so they'd better not change from page to page." (Rosenfeld et al., 2015, 147-148)	These are most commonly found at the bottom of a webpage (with links like "Sitemap," "Contact uS," "Policies," etc.	27
			Index terms	"Often referred to as keywords, tags, descriptive metadata, taxonomies, controlled vocabularies, and thesauri, sets of index term labels can be used to describe any type of content: sites, subsites, pages, content chunks, and so on." (Rosenfeld et al., 2015, 149)	"Index terms are also used to make browsing easier: the metadata from a collection of documents can serve as the source of browsable lists or menus." (Rosenfeld et al., 2015, 150)	28
		Icon labels		"These are labels that are represented by pictures or images. "We see them most frequently used as navigation system labels, especially in mobile apps where screen space is constrained." (Rosenfeld et al., 2015, 152)	These are often found on mobile websites or apps (as stated in the definition). Examples include the gear that represents setting or the three horizontal lines, which represent a navigation menu.	29
		Labeling Consistency		This refers to the grammar and structure of the labels. They need to be similar or the same in style, font, syntax, etc. to remain consistent for the users.	"Style - Haphazard usage of punctuation and case is a common problem within labeling systems, and can be addressed, if not eliminated, by using style guides. (Rosenfeld et al., 2015, 155)	30
		Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here.		31
Navigation Systems				"The use of hypertext links, icons, menu options… displayed on a Web page to move to other resources available on the Internet or to other pages within the same Web site."(ABC-CLIO, 2017, Navigation)	"Primary navigation systems that help users understand where they are and where they can go within an information environment." (Rosenfeld et al., 2015, 91)	32
	Global navigation			"…global navigation system is intended to be present on every page throughout a site. It is often implemented in the form of a navigation bar at the top of each page. These site-wide navigation systems allow direct access to key areas and functions, no matter where the user travels in the site's hierarchy." (Rosenfeld et al., 2015, 183)	Global navigation systems are the links at the top of the page or in the websites footer. For example, the links on the New York magazine – "News & Politics," "Entertainment," "Fashion," etc. never change no matter where you are in the website.	33
	Local navigation			"These local navigation systems and the content to which they provide access are often so different that these local areas are referred to as subsites, or sites within sites." (Rosenfeld et al., 2015, 187)	In The British Museum website, when you click on the "Research" (a global navigation heading) you are brought to a subsite, which has six local navigation options. Including "Collection online," "Publications," etc.	34

	Contextual navigation	"The actual definition of these links is often more editorial than architectural. Typically an author, editor, or subject matter expert will determine appropriate links once the content is placed into the architectural framework of the website. (Rosenfeld et al., 2015, 189)	Contextual navigation is using embedded links within text to move from place to place. This is essentially "contextual links" from Label Systems. See coding rules above for more information.	35
	Breadcrumb navigation	"…web browsers also support a "breadcrumbs" feature by color-coding hypertext links, a feature that can help users to retrace their steps through a website." (Rosenfeld et al., 2015, 178)	Amazon.com uses breadcrumb navigation to show where you are in the hierarchy of their site when you search for books. For example, "Kindle Store > Kindle eBooks > Computers & Technology" appears at the top of the page when viewing information architecture books.	36
	Hypertext Navigation	"Unlike physical travel, hypertextual navigation allows users to be transported right into the middle of an unfamiliar system." (Rosenfeld et al., 2015, 180)	"It is possible and often desirable to allow users to move laterally into other branches, to move vertically from one level to a higher or lower level in that same branch, or to move all the way back to the main page of the website." (Rosenfeld et al., 2015, 182)	37
	Supplemental navigation	"Supplemental navigation systems can be critical factors for ensuring usability and findability within large information systems… Supplemental navigation systems give users an emergency backup…" (Rosenfeld et al., 2015, 193) for finding information if the global, local, and contextual navigation methods don't work.	(See subcategories)	38
	Sitemaps	"A typical sitemap presents the top few levels of the information hierarchy. It provides a broad view of the content in the system and facilitates random access to segmented portions of that content via graphical or text-based links." (Rosenfeld et al., 2015, 194)	These are often located at the bottom of a page and when you click on it, it shows you the hierarchy of a site. For example, the Shoppers Drug Mart websites site map lists all the global navigation headings with the "child" webpages located beneath it.	39
	Indexes	"Similar to the back-of-book index found in many print materials, a digital index presents keywords or phrases alphabetically, without representing the hierarchy." (Rosenfeld et al., 2015, 195)	For example, the Shoppers Drug Mart website indexes "Everyday Medicines and First Aid" items in a topical index. E.g., "Eye Care," "First Aid," "Cough, Cold & Flu," etc.	40
	Guides	"Guides can take several forms, including guided tours, tutorials, and walk-throughs focused around a specific audience, topic, or task. In each case, guides supplement the existing means of navigating and understanding the system's content and functionality." (Rosenfeld et al., 2015, 198)	The w3school guides users through technology tutorials. For example, if you would like to learn more about CSS you could click on the "Learn CSS" from the homepage and select one of the many tutorials related to this subject. These will guide you through the learning process.	41
	Configurators/wizards	"Though they could be considered a special class of guide, wizards that help users to configure products or navigate complex decision trees deserve separate highlighting." (Rosenfeld et al., 2015, 200)	You use configurators or wizards when you build and price out vehicles online.	42
	Advanced Navigation	Advanced navigation approaches provide navigation specified for the individual user (in the case of personalization and customization) or through unique visualization of information.	(See subcategories)	43
	Personalization	"Personalization involves serving up information to the user based upon a	"Amazon is the most cited example of successful personalization, and some of	44

			model of the behavior, needs, or preferences of that individual." (Rosenfeld et al., 2015, 202)	the things it's done are truly valuable. It's nice that Amazon remembers our names, and it's great that it remembers our address and credit card information." (Rosenfeld et al., 2015, 202)	
		Customization	"…customization involves giving the user direct control over some combination of presentation, navigation, and content options." (Rosenfeld et al., 2015, 202)	"Customization works great for tracking the sports scores of your favorite baseball team or monitoring the value of stocks you own…" (Rosenfeld et al., 2015, 204)	45
		Visualization	"Visualization has proven most useful when the user must select among a result set of elements that she knows by their looks, as in the case of shopping for physical goods." (Rosenfeld et al., 2015, 205)	For example, shopping on Esty (the online marketplace for homemade goods) displays search results (and lists accessed by browsing) using images of the items for sale.	46
		Social Navigation	"At its simplest level, social navigation can help users discover content based on the popularity of individual items, whether by sheer volume of traffic or by implementing a user-driven voting system." (Rosenfeld et al., 2015, 206)	"Reddit, a content aggregation and discovery service, employs such a voting system—in fact, it is its primary differentiator." (Rosenfeld et al., 2015, 206)	47
	Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here.		48
Search Systems			"How search systems are powered by classification schemes, and how users use them to obtain information resources relevant to an information need/search." (The Information Architecture Institute, 2017)	"There are the guts of the search engine itself; aside from tools for indexing and spidering, there are algorithms for processing your query into something the software can understand, and for ranking the results." (Rosenfeld et al., 2015, 217)	49
	Search algorithms		"…there are algorithms for processing your query into something the software can understand…" (Rosenfeld et al., 2015, 216)	(See subcategories)	50
		Types of algorithms	Search algorithms represent the magic going on behind the scenes when you search. They scan the indexed content and return results based on your search. What is returned depends on the type of algorithm used.	"Most retrieval algorithms employ pattern matching; that is, they compare the user's query with an index of, typically, the full texts of your system's documents, looking for the same string of text. When a matching string is found, the source document is added to the retrieval set." (Rosenfeld et al., 2015, 228)	51
		Recall and Precision	Recall and precision is a mathematical calculation that determines how a search algorithm functions. Precision is calculated by dividing the number of relevant documents retrieved by the total number of document retrieved. Recall divides the total number of relevant documents retrieved by the total number of relevant documents in the system Search algorithms determine which is weighted more (recall or precision).	"Some algorithms return numerous results of varying relevance, while some return just a few high-quality results. The terms for these opposite ends of the spectrum are recall and precision." (Rosenfeld et al., 2015, 228)	52
	Search interface		"There are interfaces, too: ones for entering queries (everything from simple search boxes to advanced natural-language, voice-driven interfaces like Siri)…" (Rosenfeld et al., 2015, 216)	(See subcategories)	53
			Search is a complicated tool, but the interface shouldn't be. "… it's best to	"Consider how your search box is presented. The box can cause confusion	54

		Search interface components		keep your search interface as simple as possible: present users with a simple search box and a "search" button."" (Rosenfeld et al., 2015, 253) These are the two main components of the search interface, but they vary site to site.	when it appears alongside other boxes. Unless your system's search functionality truly requires more than one field—as is the case with many travel-related services—it is best to keep search limited to a single box." (Rosenfeld et al., 2015, 256-257)	
		Advanced search		"…advanced search interfaces allow much more manipulation of the search system and are typically used by two types of users: advanced searchers (librarians, lawyers, doctoral students, medical researchers), and frustrated searchers who need to revise their initial searches (often users who've found that the search box didn't meet their needs)." (Rosenfeld et al., 2015, 258)	Advanced search is very common on University library websites, including the University of Alberta. They allow the searcher to narrow down the results by specifying an author name, title, keyword, etc.	55
	Search results			"…you interact with the results, hopefully quickly determining which results are worth clicking through, which to ignore, and whether or not you should go back and try modifying your search." (Rosenfeld et al., 2015, 217)	(See subcategories)	56
		Displaying results		"When you are configuring the way your search engine displays results, there are two main issues to consider: which content components to display for each retrieved document, and how to list or group those results." (Rosenfeld et al., 2015, 233-234)	See subcategories)	57
			Format	This is how the search results are formatted for viewing when they are retrieved, by lists, grids, images, etc.	For example, Google lists the result, while Etsy displays search results in a grid pattern using images.	58
			Information displayed for retrieved item	This refers to what information about the search results is displayed. Is just the title displayed? Or are the title, abstract, metadata, etc. returned as well?	"Which content components you display for each result also depends on which components are available in each document (i.e., how your content is structured) and on how the content will be used. " (Rosenfeld et al., 2015, 235)	59
			Miscellaneous	This field is included so that all coded information can be categorized. Any information that does not belong in the other categories will be found here.		60
		Sorting results		This takes the search result and sorts them, alphabetically or chronologically (the two main sorting options).	(See subcategories)	61
			Alphabetically	"Just about any content component can be sorted alphabetically. This is a good general-purpose sorting approach—especially when sorting names—and in any case, it's a good bet that most users are familiar with the order of the alphabet!" (Rosenfeld et al., 2015, 240)	For example, The Baseball-Reference website sorts a search for "Donaldson" in alphabetical order. (Rosenfeld et al., 2015).	62
			Chronologically	"If your content (or your user's query) is time sensitive, chronological sorts are a useful approach. And you can often draw on a filesystem's built-in dating if you have no other sources of date information." (Rosenfeld et al., 2015, 241)	This is often used in news websites, where users are looking for the newest story first. For example, Global Edmonton presents the newest article about "Robbery" first, with the rest of the results descending by date.	63
		Ranking results		"Ranking is typically used to describe retrieved documents' relevance, from most to least. Users look to learn from those documents that are most	(See subcategories)	64

				relevant. Of course, as we shall see, relevance is relative, and you should choose relevance ranking approaches carefully. Users will generally assume that the top few results are best." (Rosenfeld et al., 2015, 240)		
			Popularity	"The popularity of the document where the query terms appear (e.g., is it linked to frequently, and are the sources of its links themselves popular?)." (Rosenfeld et al., 2015, 244)	"…Google is successful in large part because it ranks results by which ones are the most popular. It does so by factoring in how many links there are to a retrieved document." (Rosenfeld et al., 2015, 245)	65
			Users' or experts' ratings	"In an increasing number of situations, users are willing to rate the value of information. User ratings can be used as the basis of retrieval result ordering." (Rosenfeld et al., 2015, 246)	TripAdvisor returns results based on users' rating of a hotel, restaurant, or vacation activity. The highest rated item is located at the top of the page with the lowest rated at the bottom.	66
			Pay-for-placement	"Advertising has become the predominant business model for publishing online, so it is no surprise that pay-for-placement (PFP) has become commonplace in many search systems." (Rosenfeld et al., 2015, 248)	"Yelp example showed results sorted by user rankings, the first result on the list actually has a lower ranking than the others; it owes its position at the top of the list solely to the fact that it is a paid advertisement." (Rosenfeld et al., 2015, 248)	67
			Miscellaneous	This field is included so that all coded information can be categorized. Any information that does not belong in one of the three other categories will be found here.		68
		"Best bets"		"Best bets - Preferred search results that are manually coupled with a search query; editors and subject matter experts determine which queries should retrieve best bets and which documents merit best bet status." (Rosenfeld et al., 2015, 94)	While this is done on the backend of a search system. For example, when searching for news stories, the paper editors may rank human interest pieces at the top of the search results (or articles that they deem important).	69
		Hybrid (or Clustering) search results		"Hybrid approaches that combine different types of sorting—such as Google's—show a lot of promise, but you typically need to be in the business of creating search engines to have this level of involvement with a tool." (Rosenfeld et al., 2015, 248)	"Much more useful are clusters derived from manually applied metadata, like topic, audience, language, and product family. Unfortunately, approaches based on manual effort can be prohibitively expensive." (Rosenfeld et al., 2015, 248-249)	70
		Additional actions		"Contextual inquiry and task-analysis techniques will help you understand what users might want to do with their results." (Rosenfeld et al., 2015, 249)	(See subcategories)	71
			Save search	"In some cases, it's the search itself, not the results, that you're interested in "keeping." Saved searches are especially useful in dynamic domains that you'd like to track over time; you can manually re-execute a saved search on a regular basis, or schedule that query to automatically be rerun regularly." (Rosenfeld et al., 2015, 251)	The Canadian Writing Research Collaboratory (CWRC) provides logged in users with the option to save their searchers (and name them). They access the saved searches in their account dashboards.	72
			Select subset of results	"Sometimes when you're searching you want to take more than one document along with you. You want to "shop" for documents just like you shop for books at Amazon. And if you're sorting through dozens or hundreds of results, you may need a way to mark the documents you like so you don't forget or lose track of them." (Rosenfeld et al., 2015, 251)	Goodreads let users organize books by marking them "Want to Read," "Currently Reading," or "Read." If they select "Want to Read" they are saved to a list that they can return to later.	73

			Narrowing results down	“In effect, winnowing oversized result sets is a form of search revision, and often the user will self-select when he is ready to stop reviewing results. But it is still useful to provide some instruction on how to narrow search results.” (Rosenfeld et al., 2015, 264)	“A key theme in this book is the need to integrate searching and browsing (think of them together as “finding”), but we won’t belabor it here. Just remember to look for opportunities to connect your search and browse systems to allow users to easily jump back and forth.” (Rosenfeld et al., 2015, 263)	74
			Repeating/new search	“In many cases, the moment a user is confronted by a large result set is the moment he decides the number of results is too large. This is a golden opportunity to provide the user with the option of revising and narrowing his search.” (Rosenfeld et al., 2015, 238)	“If the results of a search are not satisfactory, it can be useful to state what happened behind the scenes, providing the user with a better understanding of the situation and a jumping-off point should she wish to revise her search. (Rosenfeld et al., 2015, 262)	75
		Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here		76
	Query builders			“Query builders are tools that can soup up a query’s performance. They are often invisible to users, who may not understand their value or how to use them.” (Rosenfeld et al., 2015, 229)	(See subcategories)	77
		Spell checkers		“These allow users to misspell terms and still retrieve the right results by automatically correcting search terms.” (Rosenfeld et al., 2015, 232)	“For example, “accomodation” would be treated as “accommodation,” ensuring retrieval of results that contain the correct term.” (Rosenfeld et al., 2015, 232)	78
		Phonetic tools		“Phonetic tools (the best-known of which is “Soundex”) are especially useful when searching for a name.” (Rosenfeld et al., 2015, 232)	“They can expand a query on “Smith” to include results with the term “Smyth.”” (Rosenfeld et al., 2015, 232)	79
		Stemming tools		“…a search tool might provide automatic stemming, which expands a term to include other terms that share the same root (or stem).” (Rosenfeld et al., 2015, 229)	“If the stemming mechanism is very strong, it might treat the search term “computer” as sharing the same root (“comput”) as “computers,” “computation,” “computational,” and “computing.” Strong stemming in effect expands the user’s query by searching for documents that include any of those terms. (Rosenfeld et al., 2015, 229)	80
		Natural language processing tools		“These can examine the syntactic nature of a query…”	“…for example, is it a “how to” question or a “who is” question?—and use that knowledge to narrow retrieval.” (Rosenfeld et al., 2015, 232)	81
		Autocomplete and Autosuggestions		“Autocomplete and autosuggest are widely used patterns for interacting with search systems. In both cases, a list of results is presented alongside the search box, preemptively prompting the user with possible matches based on the first few characters typed.” (Rosenfeld et al., 2015, 257)	“Displays range from very simple and straightforward text lists (in the case of autocomplete patterns) to popovers with highly customized layouts.” (Rosenfeld et al., 2015, 257)	82
	Query language			This refers to the language users use when searching (as well as what language a search system allows for.	“Further complicating the picture, there may be variations in query languages (e.g., whether or not Boolean operators like AND, OR, and NOT can be used).” (Rosenfeld et al., 2015, 216-217)	83
	Indexing content for searching			“…pieces or “atoms” of content that are typically smaller than a document. Some of that structure—say, an author’s name—may be leveraged by a search engine, while other parts—such as the legal disclaimer at the	(See subcategories)	84

				bottom of each page—might be left out.” (Rosenfeld et al., 2015, 218)		
		Search zones		“Search zones are subsets of an information environment that have been indexed separately from the rest of the content.” (Rosenfeld et al., 2015, 219)	“The creation of search zones—pockets of more homogeneous content—reduces the apples-and-oranges effect and allows users to focus their searches.” (Rosenfeld et al., 2015, 218)	85
			Navigation and destination webpages	“Most content-heavy information environments contain, at minimum, two major types of pages or screens: navigation pages and destination pages. Destination pages contain the actual information you want: sports scores, book reviews, software documentation, and so on. Navigation pages may include main pages, search pages, and pages that help you browse the environment.” (Rosenfeld et al., 2015, 220)	For example, when you click on “Research” in The British Museum you are taken to navigation page that contains links like “Collection Search,” “Research Projects” and “Blog.” When you search for the term “Blog” on the British Museum webpage this page is not in the results because it was not indexed for searching (in fact there is only one result, a link to the blog).	86
			Indexing by topic	This allows users to search and narrow search results down by topics (keywords, subjects, etc.).	“The Mayo Clinic employs topical search zones on its website. For example, if you’re looking for a doctor to help with your rehabilitation, you might select the “Doctors & Medical Staff” search zone” (Rosenfeld et al., 2015, 222)	87
			Indexing for a specific audience	These work best in audience oriented organization schemes. When a specific type of user searches, the results returned are specific to that type. Like a library search conducted by a researcher may get scholarly publications while a teen may get graphic novels.	“So we created four indexes: one for each of the three audiences, and one unified index of the entire site in case the audience-specific indexes didn’t do the trick for a particular search.” (Rosenfeld et al., 2015, 222)	88
			Indexing recent content	“Chronologically organized content allows for perhaps the easiest implementation of search zones. (Not surprisingly, it’s a common example of search zones.) Because dated materials aren’t generally ambiguous and date information is typically easy to come by, creating search zones by date—even ad hoc zones—is straightforward.” (Rosenfeld et al., 2015, 223)	When you search for publications on the University of Alberta library site you can narrow results down by publication date.	89
			Indexing full-text	“You can point your search engine at your content, tell it to index the full text of every document it finds, and let it do its thing. That’s a large part of the value of search systems—they can be comprehensive and can cover a huge amount of content quickly.” (Rosenfeld et al., 2015, 218)	“But your priority should be to set up a search system to perform full-text indexing of as much system content as possible, even across such traditional silos as company departments.” (Rosenfeld et al., 2015, 215)	90
		Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in the other category will be found here		91
	Miscellaneous			This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here		92
Vocabulary Systems				“These systems allow you to structure and map languages so that people can more easily find information.” (Rosenfeld et al., 2015, 309)	“At its simplest, a controlled vocabulary is a list of equivalent terms in the form of a synonym ring, or a list of preferred terms in the form of an authority file.” (Rosenfeld et al., 2015, 271)	93

	Metadata		"Metadata tags are used to describe documents, pages, images, software, video and audio files, and other content objects for the purposes of improved navigation and retrieval." (Rosenfeld et al., 2015, 270)	(See subcategories)	94
		Structural metadata	"Structural metadata - Describe the information hierarchy of this object." (Rosenfeld et al., 2015, 328)	Structural metadata examples include file types (JPEG, PNG, PDF), file size, and other information that directly relates to the structure of the object.	95
		Descriptive metadata	"Descriptive metadata - Think of all the different ways you might describe this object." (Rosenfeld et al., 2015, 328)	This is data that describes the object. For example, keywords, topic, etc.	96
		Administrative metadata	"Administrative metadata - Describe how this object relates to business context." (Rosenfeld et al., 2015, 328)	"Who created it? Who owns it? When was it created? When should it be removed?" (Rosenfeld et al., 2015, 328)	97
	Controlled Vocabulary		"Vocabulary control comes in many shapes and sizes. At its most vague, it consists of any defined subset of natural language. At its simplest, a controlled vocabulary is a list of equivalent terms." (Rosenfeld et al., 2015, 271)	(See subcategories)	98
		Synonym rings	"A synonym ring connects a set of words that are defined as equivalent for the purposes of retrieval." (Rosenfeld et al., 2015, 271)	"When you examine the search logs and talk with users, you're likely to find that different people looking for the same thing are entering different terms. Someone who's buying a food processor may enter "blender" or one of several product names (or their common misspellings). Take a look at the content, and you're likely to find many of these same variations." (Rosenfeld et al., 2015, 271)	99
		Authority files	"Strictly defined, an authority file is a list of preferred terms or acceptable values." (Rosenfeld et al., 2015, 275)	"The two-letter codes that constitute the standard abbreviations for U.S. states as defined by the US Postal Service provide an instructive example. Using the purist definition, the authority file includes only the acceptable codes." (Rosenfeld et al., 2015, 275)	100
		Classification schemes	"We use classification scheme to mean an arrangement of preferred terms. These days, many people prefer to use taxonomy instead. Either way, it's important to recognize that these arrangements can take different shapes and serve multiple purposes." (Rosenfeld et al., 2015, 279)	"Netflix uses a sophisticated classification scheme to help customers find new movies they may enjoy… Beyond the obvious, basic film genres ("Drama," "Comedy," etc.), Netflix movies are categorized in thousands of micro-genres, including broad ones like "Based on Real Life" and "With a Strong Female Lead," and highly specific ones like "Dark Suspenseful Gangster Dramas."" (Rosenfeld et al., 2015, 280)	101
	Thesauri		The "…thesaurus takes the form of an online database, tightly integrated with the user interface of a digital product or service. And though the traditional thesaurus helps people go from one word to many words, our thesaurus does the opposite. Its most important goal is synonym management—the mapping of many synonyms or word variants onto one preferred term or concept—so the ambiguities of language don't prevent people from finding what they need." (Rosenfeld et al., 2015, 282)	(See subcategories)	102

		Classic thesaurus		"Query terms are matched against the rich vocabulary of the thesaurus, enabling synonym management, hierarchical browsing, and associative linking. This is the full-bodied, fully integrated thesaurus we've referred to for much of this chapter." (Rosenfeld et al., 2015, 291)	"Sometimes a classic thesaurus isn't practical because of issues on the content side of the equation that prevent document-level indexing." (Rosenfeld et al., 2015, 292)	103
		Indexing thesaurus		"It allows you to build browsable indexes of preferred terms, enabling users to find all documents about a particular subject or product through a single point of access." (Rosenfeld et al., 2015, 291)	"It structures the indexing process, promoting consistency and efficiency. The indexers can work as an integrated unit, given a shared understanding of preferred terms and indexing guidelines." (Rosenfeld et al., 2015, 291)	104
		Searching thesaurus		"A searching thesaurus leverages a controlled vocabulary at the point of searching but not at the point of indexing." (Rosenfeld et al., 2015, 291)	"You also have the option of giving more power and control to the users—asking them whether they'd like to use any combination of preferred, variant, broader, narrower, or associative terms in their queries. When integrated carefully into the search interface and search result screens, this can effectively arm users with the ability to narrow, broaden, and adjust their searches as needed." (Rosenfeld et al., 2015, 292)	105
		Thesaurus standards		"The standard provides a valuable conceptual framework and in some cases offers specific rules you can follow, but it absolutely does not remove the need for critical thinking, creativity, and risk taking in the process of thesaurus construction." (Rosenfeld et al., 2015, 293)	"The ANSI/NISO standard is entitled "Guidelines for the Construction, Format and Management of Monolingual Thesauri." (Rosenfeld et al., 2015, 293) There are hundreds of thesaurus standards and this is just one example.	106
		Thesaurus terms		"If you're working with… thesauri, it's useful to know the core terminology used by experts in the field to communicate definitions and relationships. This specialized technical language can provide efficiency and specificity when communicating among experts." (Rosenfeld et al., 2015, 283)	"Preferred term (PT) Also known as the accepted term, acceptable value, subject heading, or descriptor. All relationships are defined with respect to the Preferred Term." (Rosenfeld et al., 2015, 283)	107
		Miscellaneous		This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here		108
	Semantic relationships			This is the relationship between indexed terms in your site (defined according to your thesaurus).	(See subcategories)	109
			Hierarchical relationships	"The hierarchical relationship divides up the information space into categories and subcategories, relating broader and narrower concepts through the familiar parent–child relationship. (Rosenfeld et al., 2015, 296)	"At first blush, the hierarchical relationship sounds pretty straightforward. However, anyone who's ever developed a hierarchy knows that it isn't as easy as it sounds. There are many different ways to hierarchically organize any given information space (e.g., by subject, by product category, or by geography)." (Rosenfeld et al., 2015, 297)	110
			Polyhierarchical relationships	"Or, if you're pragmatic, you can allow for some level of polyhierarchy, permitting some terms to be cross-listed in multiple categories. When you're dealing with large information systems, polyhierarchy is unavoidable. As the number of documents grows, you need a greater	"In digital information systems, the only real challenge introduced by polyhierarchy is representing the navigational context. Most systems allow for the notion of primary and secondary locations within the hierarchy." (Rosenfeld et al., 2015, 303)	111

				level of precoordination (using compound terms) to increase precision, which forces polyhierarchy." (Rosenfeld et al., 2015, 302)		
			Equivalence relationships	"The equivalence relationship is employed to connect preferred terms and their variants. While we may loosely refer to this as "synonym management," it's important to recognize that equivalence is a broader term than synonymy." (Rosenfeld et al., 2015, 295-296)	"Depending on the desired specificity of your controlled vocabulary, you may also fold more general and more specific terms into the equivalence relationship to avoid extra levels of hierarchy." (Rosenfeld et al., 2015, 296)	112
			Associative relationships	"Associative relationships allow what marketing folks call "cross-selling," allowing an ecommerce site, for example, to say "Hey, nice trousers! They'd go great with this shirt."" (Rosenfeld et al., 2015, 298)	"There is the notion that associative relationships should be "strongly implied." For example, Hammer RT Nail. In practice, however, defining these relationships is a highly subjective process." (Rosenfeld et al., 2015, 298)	113
	Faceted classification			"A classification system developed through analysis of the fundamental characteristics of subjects by which they can be divided into subclasses. For example, in his Colon Classification, S.R. Ranganathan identifies five basic characteristics: personality, matter, energy, space, and time (abbreviated PMEST). In such a system, the notation representing a subject is created by combining the notations of its facets." (ABC-CLIO, 2017, Faceted Classification)	"The designers of Wine.com have made decisions throughout the site about how and when to leverage facets within the interface. For example, you can browse by ratings from individual magazines from the main page. Hopefully, these are informed decisions made by balancing an understanding of user needs (how people want to browse and search), business needs (how Wine.com can maximize sales of high-margin items), and the creation of meaningful contexts." (Rosenfeld et al., 2015, 307)	114
	Miscellaneous			This field is included so that all coded information can be categorized. Any information that does not belong in one of the other categories will be found here		115

Most definitions and examples found in this code frame were taken from the pilot study.

The creation of the coding frame relied on instructions/information found in *Qualitative content analysis in practice* by Margrit Schreier.

Sources for Descriptions:

ABC-CLIO (2017). Navigation. Retrieved from: http://www.abc-clio.com/

Morville, P., & Rosenfeld, L. (2006) *Information Architecture for the World Wide Web (3rd ed)*, Sebastopol, CA: O'Reilly Media Inc. http://proquest.safaribooksonline.com.login.ezproxy.library.ualberta.ca/book/web-development/0596527349/firstchapter.

Rosenfeld, L., Morville, P. & Arango, J. (2015). *Information architecture: For the web and beyond* (4th ed.). Sebastopol, CA: O'Reilly Media Inc.

Schreier, M. (2012). *Qualitative content analysis in practice*. London, UK: Sage Publications Ltd.

The Information Architecture Institute (2017). What is Information Architecture? Retrieved from: http://www.iainstitute.org/what-is-ia Accessed April 25, 2017.

Appendix H – Content Analysis Form

Source ID	Unit ID	Text excerpt	Text Summary	Code	Main Category	Subcategory 1	Subcategory 2	Subcategory 3	Researcher notes

Appendix I – Content Analysis Statistics

Name	Frequency
Organization system	585
Labeling system	279
Navigation system	636
Search system	450
Vocabulary system	566
Total	2516

Name	Percentage
Organization system	23.25%
Labeling system	11.09%
Navigation system	25.28%
Search system	17.89%
Vocabulary system	22.50%
Total	100.00%

The statistics were created in Microsoft Excel 2010.

- Frequency formula =COUNTIF(range, criteria).
  - Example: =COUNIF(F581:F964, G2560). “F581:F964” is the range that all of source 2's coded text falls between and “G2560” is the criteria (the term “Labeling System), so this determines how many coded excerpts are classified under “Labeling system” within source 2.
- Percentage formula =Number/$Column$Number
  - Example: Calculating the frequency of Navigation systems would look like this: =H2528/H2531

Source	Frequency	Percentage
1	579	23.01%
2	384	15.26%
3	1481	58.86%
4	72	2.86%
Total	2516	100.00%

Name	Frequency: Source 1	Percentage
Organization system	53	9.15%
Labeling system	40	6.91%
Navigation system	233	40.24%
Search system	95	16.41%
Vocabulary system	158	27.29%
Total	579	100.00%

Name	Frequency: Source 2	Percentage
Organization system	200	52.08%
Labeling system	39	10.16%
Navigation system	119	30.99%
Search system	2	0.52%
Vocabulary system	24	6.25%
Total	384	100.00%

Name	Frequency: Source 3	Percentage
Organization system	287	19.38%
Labeling system	194	13.10%
Navigation system	282	19.04%
Search system	352	23.77%
Vocabulary system	366	24.71%
Total	1481	100.00%

Name	Frequency: Source 4	Percentage
Organization system	45	62.50%
Labeling system	6	8.33%
Navigation system	2	2.78%
Search system	1	1.39%
Vocabulary system	18	25.00%
Total	72	100.00%

Miscellaneous Categories		
Code	Frequency	Percentage
13	14	13.59%
21	7	6.80%
31	5	4.85%
48	47	45.63%
60	2	1.94%
68	12	11.65%

76	5	4.85%
91	1	0.97%
92	8	7.77%
108	1	0.97%
115	1	0.97%
Total	103	100.00%

Total Miscellaneous in each category		
Name	Frequency	Percentage
Organization system	21	20.39%
Labeling system	5	4.85%
Navigation system	47	45.63%
Search system	28	27.18%
Vocabulary system	2	1.94%
Total	103	100.00%

Miscellaneous percentage of total
4.09%

Appendix J – Coded Data Analysis for Heuristics

	3	1264	As a result, social classification—primarily driven by user-generated content tagging—has emerged as an important tool for organizing information in shared information environments. P. 127	Social classification (Social tagging) is primarily user driven (organization structure)	19	Organization system	Organization structure	Social Tagging		Definition
	3	1265	Free tagging, also known as collaborative categorization, mob indexing, and ethnoclassification, is a simple yet powerful tool. Users tag objects with one or more keywords. P. 127	Social Tagging is a powerful tool to organize information	19	Organization system	Organization structure	Social Tagging		
	3	1266	These tags can be informally supported in text fields, or they can be provided for with bespoke fields in the formal structure of content objects. P. 127	Tags can be informal or formal	19	Organization system	Organization structure	Social Tagging		Definition
	3	1268	No single person or centralized team created a taxonomy to define these relationships. Rather, they emerged (and continue to emerge) through the tagging efforts of many individuals. P. 127-128	Tagging organization emerges through the tagging by individuals	19	Organization system	Organization structure	Social Tagging		Definition
	3	1269	Similarly, LinkedIn allows users to "endorse" their professional contacts as possessing certain individual professional skills (Figure 6-20). These endorsements are in effect tags: they allow users to describe their business contacts in a granular way that informs how the system groups them with similar people. p. 128	LinkedIn uses social tagging to endorse their professional contacts	19	Organization system	Organization structure	Social Tagging		Example
	3	1270	an impassioned debate raged over whether or not free-form tag structures (or "folksonomies," as information architect Thomas Vander Wal cleverly christened them) would eliminate the need for top-down, centrally defined information structures. P. 128	While social tagging structure are new, they will not replace tried and true methods	19	Organization system	Organization structure	Social Tagging		
	3	1272	high-profile experiments in tag-driven systems—such as the bookmarking service Delicious.com—fizzled in the marketplace, and most of these systems employed tags within centrally defined structures anyway. p. 128	Tag-driven structures like Delicious.com have fizzled	19	Organization system	Organization structure	Social Tagging		
	3	1273	Still, free-form tagging has proven its usefulness in specific situations, and it remains a valuable tool in the information architect's toolset. P. 128-129	Social tagging has proven useful	19	Organization system	Organization structure	Social Tagging		
	3	1298	or user-contributed tagging. P. 130	Social tagging structure	19	Organization system	Organization structure	Social Tagging		
	3	1314	Social classification has emerged as an important tool for organizing information in shared digital environments. P. 131	Social classification (Social tagging) for organizing info	19	Organization system	Organization structure	Social Tagging		
	2	602	When you create a set of categories and assign your content to it, some content will definitely fit in more than one place (though if it happens a lot, it may be a sign you need to think more about your approach). P. 115	content fits into more than one type of category (reexamine scheme if it happens too often)	20	Organization system	Organization structure	Hybrid structure		Definition
	2	753	Many sites that start with a hypertext structure are later re-organised, when the content is known. P. 188	Sites may start with hypertext then later reorganize	20	Organization system	Organization structure	Hybrid structure		
	2	760	Now let's look at how these three simple patterns can be combined to create more complex information architectures. P. 189	These structures can be combined	20	Organization system	Organization structure	Hybrid structure		
	2	761	One very common pattern (probably the most common pattern on the web) is a combination of a simple hierarchy with some database content. P. 191	Hierarchy combined with database structure	20	Organization system	Organization structure	Hybrid structure		Type #1
	2	762	This pattern is suitable for all sorts of small, medium and large sites. It lets you create hierarchical sections of the website for basic content, and then to use the power of a database to assemble detailed information within a section. P. 191	Hierarchy and database structure means a hierarchy access to the content and then content structures in the database model	20	Organization system	Organization structure	Hybrid structure		
	2	763	The hierarchy and database pieces may be integrated (e.g. a database structure within one of the sections of the site), side by side (e.g. a database structure as a whole section of the site) or any combination in between. P. 191	Hierarchy and database can be combined in just one section, through whole site or in between	20	Organization system	Organization structure	Hybrid structure		
	2	764	One of the main challenges with this type of pattern is deciding what pieces you'll turn into structured content and what you'll leave as hierarchical. P. 192	Hierarchy-database structure problem is deciding what you will structure and what you leave as hierarchy	20	Organization system	Organization structure	Hybrid structure		
	2	765	Database structures can help you manage larger volumes of information. If you have a handful of news stories a year, you can manage them as hierarchical content. If you have hundreds a day, you'll probably want to leverage the power of the database structure to automate the display. p. 193	Database structures help manages large volumes of content - hierarchies smaller collections	20	Organization system	Organization structure	Hybrid structure		

2	821	Make sure your labels clearly explain the content they represent. P. 213	Make sure the explain what they represent	30	Labeling systems	Labeling consistency				
2	822	My daughter's school's website has a section called Parents (not For Parents, just Parents). I thought it might contain news specific for parents, or ways for parents to be involved in the school (after all, it's a school website – most of the content is for parents). Imagine my surprise when I found out this is where the canteen and after-school care information was – something I'd previously looked for and thought wasn't around. The label just doesn't communicate what is included in the section. P. 213	Example of unclear labeling	30	Labeling systems	Labeling consistency				Example
2	823	Don't use cute, obscure labels. Well, I guess you can, but you risk people not clicking on them. P. 213	Don't use "cute" labels unless you want to risk users not clicking on them	30	Labeling systems	Labeling consistency				
2	824	If you absolutely know your audience is willing to explore your site, go ahead. But if yours isn't the type of site people will be willing to explore, make your labels plain and clear. P. 213	If you know your audience is willing to explore site, use niche labeling	30	Labeling systems	Labeling consistency				
2	825	Clear labels aren't necessarily formal. Be consistent with your brand and writing style – use friendly labels if that's your voice, or use more formal labels if your writing is more formal. P. 213	Clear labels don't have to be formal	30	Labeling systems	Labeling consistency				Component
2	826	Content: look at the words used in your content and the labels they suggest. P. 214	Look at your content to find labels	30	Labeling systems	Labeling consistency				
2	827	User research: your research will contain hundreds of words that describe how your audience phrase ideas and concepts. Make sure you look at your search logs – you'll see exactly what people are looking for. P. 214	User research to create labels	30	Labeling systems	Labeling consistency				
2	828	Card sorting: in a card sort, the last step is often for people to provide a label describing what the group of cards is about. I often find these labels a bit long or informal to use, but they can give you some ideas. P. 214	Use card sorting exercise to create labels	30	Labeling systems	Labeling consistency				
2	829	What everyone else does: look around at competitors or other content like yours and see what other people do. If there's a consistent approach, you can do your users a favour and use that. If you're tossing up between two terms, use google trends to see the popularity of particular terms. P. 214	Look at competitors to see what they do - don't have to copy but it helps users feel comfortable in websites	30	Labeling systems	Labeling consistency				
2	830	Finally, expect labels to change over time. As you add more content to a site, or move things around, labels will need to change as well. Language is subtle and slight changes in words may be needed over time. P. 214	Expect labels to change over time - keep them updates	30	Labeling systems	Labeling consistency				Component
2	832	Look at your draft groups. Think about what you know about your audience. (Start by thinking about just your core audience or core tasks.) P. 218	look at categories and think about audience to create labels	30	Labeling systems	Labeling consistency				
2	833	If you're working with database content: Look at the attributes of each type of content. Look at what you know about your audience and what they need. P. 219	If working with database content look at attributes for content	30	Labeling systems	Labeling consistency				
3	1002	Labeling systems describe categories, options, and links in language that (hopefully) is meaningful to users; you'll see examples throughout the page (e.g., "Admission," "Alumni," "Events"). P. 83	Labeling system uses language meaningful to the user	30	Labeling systems	Labeling consistency				Definition
3	1025	How we represent information—for example, using scientific terminology ("Acer") or lay terminology ("maple") p. 90	How to use labels - consistency (choose the right form)	30	Labeling systems	Labeling consistency				
3	1332	While this label doesn't seem to be that ambiguous in a desktop browser—after all, the rest of the site menu items are laid out next to it—it becomes more problematic when rendered in a mobile browser, because the label "Menu" is more often experienced in those browsers as a way of accessing the system's main navigation menu p. 136-137	Labeling Menu is ambiguous, especially is it's used in a cell phone (it would then have another menu label)	30	Labeling system	Labeling consistency				
3	1334	It gives the impression that "Coffeehouse" has a particular meaning within the Starbucks Corporation, and one that is not immediately evident from examining the content it represents. P. 138	Ambiguous labeling that is unclear for users	30	Labeling system	Labeling consistency				Example
3	1335	It's also worth noting that because "Coffeehouse" begins with the word "Coffee" (and sits close to it in the navigation menu), this label may cause users to do a double-take when looking for coffee. P. 138	Having two "Coffee" based navigation labels confusing	30	Labeling system	Labeling consistency				
3	1338	confusion for users who may read it as leading to information	confusing	30	Labeling system	Labeling consistency				

3	1333	the desktop version of the Starbucks website reveals mega-menus for each label, mobile users can't derive additional contextual clues; all they have to go by are the labels of the global navigation menu. p. 137	menus from its global navigation bar to give more info (but mobile users don't know that)	33	Navigation System	Global Navigation	
3	1337	This is the only verb in the global navigation, a potential source of confusion for users who may read it as leading to information about a physical "shop." p. 139					
3	1520	Integrating global, p. 175	Global navigation of Starbucks	33	Navigation System	Global Navigation	
3	1533	First, we have the global, . P. 176	Global navigation	33	Navigation System	Global Navigation	
			Global navigation	33	Navigation System	Global Navigation	
3	1537	The need for global, P. 176	Global navigation in mobile environments	33	Navigation System	Global Navigation	
3	1552	As soon as we start talking about global, p. 178	Global navigation	33	Navigation System	Global Navigation	
3	1555	Does the local navigation bar work best at the top of the page, or is it better running down the left side? P. 178	Where should the global navigation go	33	Navigation System	Global Navigation	
3	1556	Should we use mega-menus or fat footers to reduce the required number of clicks? P. 178	Mega menus and fat footers reduce the number of clicks	33	Navigation System	Global Navigation	
3	1567	Global, P. 183	Global navigation	33	Navigation System	Global Navigation	Component
3	1570	By definition, a global navigation system is intended to be present on every page throughout a site. It is often implemented in the form of a navigation bar at the top of each page. P. 183	Global navigation is present on every page of the site (often as navigation bar)	33	Navigation System	Global Navigation	Definition
3	1571	These site-wide navigation systems allow direct access to key areas and functions, no matter where the user travels in the site's hierarchy. P. 183	Global navigation allows direct access to key areas of the site anywhere in the site	33	Navigation System	Global Navigation	Definition
3	1572	Global navigation bars come in all shapes and sizes. P. 184	Global navigation design varies	33	Navigation System	Global Navigation	
3	1573	Most global navigation bars provide a link to the home page, usually represented as the organization's logo. P. 184	Global navigation can be found in the logo related to the site	33	Navigation System	Global Navigation	
3	1574	Some, like Apple's and Acer's, reinforce the site's structure and provide contextual clues to identify the user's current location within the site. Others, like Dell's, have a simpler implementation and don't do either. P. 184	Global navigation can identify the users current location or not - it varies site to site	33	Navigation System	Global Navigation	Example
3	1575	Global navigation system design forces difficult decisions that must be informed by user needs and by the organization's goals, content, technology, and culture. One size does not fit all. P. 184	Global navigation design depends on your site, users, content, context, etc.	33	Navigation System	Global Navigation	
3	1576	Global navigation bars are constantly evolving. For example, in recent years mega-menus and fat footers have become common design patterns for rendering global navigation structures in websites. P. 184	Global navigation design is evolving - development of fat footers and mega menus	33	Navigation System	Global Navigation	Components
3	1577	Mega-menus are like traditional drop-down menus: usually rendered at the top of a page, they provide access to second- and third-level elements when the user clicks on a first-level element. However, mega-menus are much richer than the simple lists of links of yesteryear; they often feature sophisticated typographic layouts, images, and other cues to give the user insight into the content and structure of the system p. 184	Mega menus are like traditional drop down menus, but they provide access to second and third level pages, and are sophistically designs	33	Navigation System	Global Navigation	
3	1578	Fat footers are abridged... rendered at the bottom of web pages. They provide direct access to the most important sections of the site p. 185	Fat footers are at the bottom of the page - most important sections of the site	33	Navigation System	Global Navigation	
3	1580	Because global navigation bars are often the single consistent navigation element in the site, they have a huge impact on usability. P. 186	Global navigation are on every page so usability is key	33	Navigation System	Global Navigation	Pros/Definition
3	1581	On many websites, the global navigation system is complemented by p. 186	Global navigation is often complimented by	33	Navigation System	Global Navigation	
3	1583	Some tightly controlled sites integrate global p. 186	Global navigation	33	Navigation System	Global Navigation	
3	1585	For example, the USA Today website presents a global navigation bar that shows p. 186	USA today website presents global navigation	33	Navigation System	Global Navigation	
3	1587	A reader who selects "Money" sees p. 186	Global navigation	33	Navigation System	Global Navigation	
3	1589	selects "Life," but both sets of options are presented within the same navigational framework p. 186	Global navigation category in USA today	33	Navigation System	Global Navigation	
3	1591	or with the global navigation system. P. 187	Global navigation	33	Navigation System	Global Navigation	
3	1599	make creaeting global navigation systems look easy. P. 188	Global navigation look easy	33	Navigation System	Global Navigation	
3	1600	Some relationships don't fit neatly into the structured categories of global P. 188	Not everything fits within global navigation	33	Navigation System	Global Navigation	
3	1620	One key to success is simply recognizing that global, p. 191	Global navigation	33	Navigation System	Global Navigation	
3	1624	In some cases, you may need to revisit the number of options within each navigation bar. P. 191	May need to limit global navigation bar options	33	Navigation System	Global Navigation	

3	1921	showing the query terms within the "context" of the document's text is a useful variation on this theme p. 236	Show the query term highlighted in the information returned	59	Search systems	Search results	Displaying results	Information displayed for retrieved results	
3	1922	The Verge highlights the query terms by using a bold font within the sentence they appear in—an excellent practice, as it helps the user quickly scan the results page for the relevant part of each result. p. 236	The verge highlights the query terms (bolded) in the sentence that they appear in	59	Search systems	Search results	Displaying results	Information displayed for retrieved results	Example
3	1924	If your engine is configured to display a lot of information for each retrieved document, p. 238	if your engine is configured to display lots of info	59	Search systems	Search results	Displaying results	Information displayed for retrieved results	
3	988	For example, how search results are presented can suggest different "beats," with some environments requiring denser patterns than others p. 61	Search result presentation can differ depending on the rhythm chosen by designers	60	Search systems	Search results	Displaying results	Miscellaneous	
3	1927	We suggest that you let users know the total number of retrieved documents so they have a sense of how many documents remain as they sift through search results. P. 238	Let users know how many results were returned	60	Search systems	Search results	Displaying results	Miscellaneous	
3	1058	sorted, 93	How search results are sorted	61	Search systems	Search results	Sorting results		
3	1253	sorting of search results p. 125	Sorting search results	61	Search system	Search results	Sorting results		
3	1906	and how to list P. 234	How to list results	61	Search systems	Search results	Sorting results		
3	1934	There are two common methods for listing retrieval results: sorting p. 240	Sorting search results	61	Search systems	Search results	Sorting results		Definition
3	1941	Sorting is especially helpful to users who are looking to make a decision or take an action. For example, users who are comparing a list of products might want to sort by price or another feature to help them make their choice. P. 240	Sorting is helpful for users who are looking to make decision	61	Search systems	Search results	Sorting results		〃
3	1942	Any content component can be used for sorting, but it's sensible to provide users with the option to sort on components that will actually help them accomplish tasks. P. 240	Best to sort results using component that will help them accomplish tasks	61	Search systems	Search results	Sorting results		〃
3	1980	one alternative approach to sorting p. 248	Alternative to sorting	61	Search systems	Search results	Sorting results		
3	2055	Showing other current settings, such as the sort order p. 262	Show sort order	61	Search systems	Search results	Sorting results		
4	2469	Sorting is the act of arranging content according to established rules. P. 103	Sorting is the act of arranging with rules	61	Search systems	Search results	Sorting results		〃
3	1937	or alphabetically by any number of content component types (e.g., by title, by author, or by department). P. 240	Results sorted alphabetically	62	Search systems	Search results	Sorting results	Alphabetically	Definition
3	1946	Just about any content component can be sorted alphabetically p. 240	Any results can be sorted alphabetically	62	Search systems	Search results	Sorting results	Alphabetically	Definition
3	1947	This is a good general-purpose sorting approach—especially when sorting names—and in any case, it's a good bet that most users are familiar with the order of the alphabet! P. 240	Everyone understands the alphabet, so this is an easy way to sort results	62	Search systems	Search results	Sorting results	Alphabetically	Definition
3	1948	It works best to omit initial articles such as "a" and "the" from the sort order (certain search engines provide this option); users are likely to look for "The Naked Bungee Jumping Guide" under "N" rather than "T." p. 240	Best to remove the articles (a, the, etc.) from the sorting (people wont look in "The" first)	62	Search systems	Search results	Sorting results	Alphabetically	Example
3	1936	Retrieval results can be sorted chronologically by date, p. 240	Results sorted chronologically	63	Search systems	Search results	Sorting results	Chronologically	Definition
3	1949	If your content (or your user's query) is time sensitive, chronological sorts are a useful approach. P. 241	If the material returned is time sensitive you can sort by date	63	Search systems	Search results	Sorting results	Chronologically	〃
3	1950	And you can often draw on a filesystem's built-in dating if you have no other sources of date information. P. 241	When sorting by date you can draw on built-in dating info	63	Search systems	Search results	Sorting results	Chronologically	〃
3	1951	If your site provides access to press releases or other news-oriented information, sorting by reverse chronological order makes good sense (see Figure 9-17 and Figure 9-18). P. 241-242	Sorting by date makes sense in news oriented sites	63	Search systems	Search results	Sorting results	Chronologically	〃
3	1952	Chronological order is less common and can be useful for presenting historical data. P. 242	Chronological order is useful for historical data	63	Search systems	Search results	Sorting results	Chronologically	〃
1	267	In the beginning, you can use subject matter experts (SME) to rank your results in terms of relevance. P. 103	Initially you can use SMEs to rank results in terms of relevance	64	Search system	Search results	Ranking results		
3	1057	and how sets of results should be ranked, p. 93	How search results are ranked	64	Search systems	Search results	Ranking results		
3	1782	and for ranking the results. P. 216	Ranking search results	64	Search systems	Search results	Ranking results		
3	1907	or group those results. P. 234	How to group results	64	Search systems	Search results	Ranking results		
3	1933	what order should these results be listed? Again, much of the answer depends upon what kind of information needs your users start with, what sort of results they are hoping to receive, and how they would like to use the results. P. 239-240	The order used to present search results depends on the user	64	Search systems	Search results	Ranking results		Definition
3	1935	and ranking. P. 240	ranking search results	64	Search systems	Search results	Ranking results		

1	135	A different image of excess (descriptive) p. 73	You can search using descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
1	143	But add some descriptive metadata: "San Francisco Giants, home run, World Series chances, Barry Bonds, baseball." And now a search will quite likely give you this column in its results.	Descriptive metadata about an article makes it easier to find	96	Vocabulary systems	Metadata	Descriptive metadata		
1	225	While the descriptive metadata falls into the messy "tagging" bucket, p.91	Etsy uses descriptive metadata (the tags)	96	Vocabulary systems	Metadata	Descriptive metadata		Definition
1	227	It's still tagging, in that the users of the site add the keywords, not the administrators of the site. P. 91	Descriptive metadata is still tagging in Etsy because it is done by users	96	Vocabulary systems	Metadata	Descriptive metadata		Example
1	558	and descriptive.	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
1	562	and Topic and Level are both descriptive. P. 261	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
1	566	Right now, Boxes and Arrows is only using Category metadata. P. 262	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
1	568	When we sat down to talk with the engineers, they mentioned keywording (adding keywords) was trivial to implement, as long as it was just another field in the story's metadata area. P. 263	Adding keywords field to metadata area - descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
2	739	Descriptive: Description of the item P. 186	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
2	742	Title (descriptive) Category (descriptive) Tags (descriptive) P. 186-187	Descriptive metadata (examples)	96	Vocabulary systems	Metadata	Descriptive metadata		Examples
3	1424	descriptive metadata, P. 149	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
3	2342	With the underlying descriptive metadata P. 307	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
3	2373	Descriptive metadata Think of all the different ways you might describe this object. How about topic, audience, and format? There should be at least a dozen different ways to describe many of the objects you study. Now is the time to get them all on the table. P. 328	Descriptive metadata tags keywords and information that describes the objet	96	Vocabulary systems	Metadata	Descriptive metadata		Definition
3	2388	and descriptive metadata fields. P. 375	Descriptive metadata	96	Vocabulary systems	Metadata	Descriptive metadata		
3	2409	Descriptive metadata needs to be created to more easily populate the site architecture with relevant content. For instance, each news story blurb on the "Weather in the News" main page should make use of the descriptive data shown	Descriptive metadata needs to be created to populate information architecture	96	Vocabulary systems	Metadata	Descriptive metadata		Definition
1	108	Administrative. Metadata about the way the thing will be handled. Is it a temporary thing, or does it need to be archived? Who is the editor? Has it been approved for publication? P. 66	Definition of Administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		Definition
1	111	Administrative. Because that's what the item's purpose is. P. 67	Is "Christmas card" administrative metadata?	97	Vocabulary systems	Metadata	Administrative metadata		
1	115	And on HomeBistro.com, we see this administrative metadata: <meta name="ROBOTS" content="ALL"> <meta name="revisit" content="15 days"> <meta name="robots" content="index, follow"> p. 67	HTML meta tags showing Administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		Example
1	123	You might remember that it was taken by someone named Beatrice Abbott. Or that it was taken in 1936. This is administrative metadata and includes not only the author/-creator of the information, but the date created, the date published, and so on—everything about how the item/information was managed.	Examples of administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		Example
1	130	A user could search for "Bernice Abbott," the photographer, or Changing New York, the series name. p. 70	Find items using administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		
1	136	A different photo by the photographer (administrative) p. 73	You can search using administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		
1	496	Hulu uses administrative metadata to determine recently added videos, p. 212	Hulu uses administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		Example
1	557	administrative,	Administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		
1	561	Category is administrative, p. 261	Administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		
2	738	Administrative: How it is used P. 186	Administrative metadata	97	Vocabulary systems	Metadata	Administrative metadata		
2	741	Author (administrative) Date posted (administrative) URL (administrative) Status: published (administrative) P. 186	Administrative metadata (example)	97	Vocabulary systems	Metadata	Administrative metadata		Example

Appendix K – Heuristic Definitions

Organization Systems:

Organization systems shape and present information in a variety of ways, and we use these systems to make sense of website information (998, 2468). And the creation of these systems serve as a set of instructions for people interacting with a website (2486), the more familiar people are with organization system, the easier it is for them to use (46, 55).

Organization schemes:

Organization schemes define the shared characteristics of content and determine how they are grouped together (1098).

Example: We navigate through organization schemes everyday – in real life and online. These include contact directories, supermarkets, and libraries. They all use organization schemes to organize and facilitate information access (1105).

Exact organization scheme

This classification of schemes divides the website organization into well-defined and mutually exclusive sections (1109).

Pros and cons:

- Easy to design and maintain the separate categories (1112).
- Best for known item searching, users need to know the specific name of the item they are searching for or it will slow the users down (1290, 1833, 1111).
- You can lose flexibility of organization the more exact your scheme is, a problem if you introduce an item that doesn't fit into one of the categories (2477).
- Exact organization schemes can slow you down – different people will classify differently (e.g., a tomato can be classified as a fruit, veggie, etc.) (2478, 2484)

Alphabetical scheme:

Alphabetical schemes organize information by the alphabet (an exact scheme), as long as something has a "title" it can be organized alphabetically (638, 639).

Examples:

- Encyclopaedias and dictionaries (113, 642, 2445).
- Great for A-Z indexes (639, 840)

Cons:

- The grouped items generally have nothing in common other than the first letter of their name (some exceptions, e.g., country names) (1133).

Chronological schemes:

This scheme organizes events/items by dates, which is possible when content has time as a key aspect like historical events or news. As long as there is agreement on when they occurred chronological schemes are easy to design and use (492, 627, 628, 1121).

Examples:

- News weblogs, history, TV guides, event listings (628)
- Press releases (1118)
- History books, magazine archives, diaries, TV guides. (1120)

Components:

- You need to think about how to order the list (oldest to newest or vice versa) (637).
- You can use the chronological scheme to group or sequencing method (636).

Geographical scheme:

Used to organize information by location (if they have location as a key attribute). These are very straightforward to design; they either have a location or do not (though sometimes there are issues with border disputes). (650, 1124)

Examples:

- Craig's List has users select their location before searching for items (1125).
- Maps for displaying the organized information (but when using these figure out exactly what you want to achieve with this, e.g., trying to show where items are located exactly in the map or in a particular area) (665, 657, 660).
- News or weather that affects a local area (1122).

Components:

- Your audience must want to access information geographically (654).
- Your users must understand the geography you're using (often in a lot of detail) (654).

Format scheme (from the organization scheme miscellaneous category):

Format is an exact organization scheme. You organize content by the file format. You can organize by format "types" like videos, articles, etc. or you can organize the files by structural metadata (e.g., JPEG, PNG, etc.) (664, 1089).

Examples:

- Instructional websites (where they group videos, articles and tutorials) and article websites (where they group articles, interviews, and tools) (664).

Organizational (business) organization scheme (from the organization scheme miscellaneous category):

These are based on the unique organization of businesses. This works the best for company intranets. This scheme works well when each business department takes control of (and the responsibility for) their sections of their website (667, 668, 669, 670, 1820).

Example:

- Defense department, the author kept the organizational organization scheme for the intranet (671).

Ambiguous organization schemes:

These divide information into categories that defy exact definitions. They are ambiguous in language, organization, and human subjectivity (1126). This type of organization scheme requires more thought about how to classify information, the more ambiguous, the more it can be argued about (2472).

Pros and cons:

- Ambiguous organization schemes are best for browsing the information (you don't have to know exactly what you're looking for) (1291).
- These schemes are difficult to update and maintain (1138).

Topical organization scheme:

Topical organization schemes organize information by topic/subject (or what they're about). This works for almost every website and is the most commonly found organization scheme (605, 691, 693, 692, 1023).

Examples:

- Hulu – TV, movies, most popular, recently added (495, 498).
- Newspapers – local, world, business, etc. (53, 1140).
- Australian science site – Space, Agriculture, Technology, etc. (710).
- Consumer reports website relies on topical organization scheme (1145).

Components:

- You need to define the breadth of topical organization scheme coverage. E.g., Encyclopaedias cover all knowledge, but websites cover local information (114, 1146).

Task oriented organization scheme:

This scheme organizes information around the main task people perform on your website. This is the best choice when there are only small amount of tasks and when you can anticipate the tasks that the users need to perform on your site (674, 677, 1149).

Examples:

- Customer interaction websites (115).
- Desktop and mobile apps that support the creation and management of content (word processors, etc.) (1150).
- Intranets and extranets lend themselves to task organization scheme because they integrate application with content (1152)

Components:

- Task organization needs to have clear boundaries, they can't overlap (675).
- When classifying task organization look for phrases like "I need" or "I do," whatever follows is usually a task. (678)

Audience specific organization scheme:

These schemes organize information by the different type of audience or user that use your site. This organization scheme breaks the website in to smaller sites (that are audience specific), allowing a clutter free design that presents only the options of interest to a particular audience (1156).

Examples:

- Nordstrom and the Gap organize clothes into "Men," "Women," and "Kids" (38, 39).

- CERN present audience oriented organization scheme using the terms like "Scientists," to get users to self identify (1157, 1159).

Components:

- Need to be aware of different audience levels (beginner, intermediate, expert, etc.) (573).
- Audience organization scheme need to have very clear boundaries between audiences. Duplication of information between audience areas can be confusing and hard to maintain (682, 687, 704).
- Audience specific schemes can be either open or closed. Open means that any audience type can access every area of a website. Closed means that you can only access information in your audience category (1161, 1162, 1163).

Metaphor driven organization scheme:

This type of organization scheme organizes a website around real world environments, relating the site to the familiar (1164). There are three different types of metaphor driven schemes:

Examples:

- Shopping for groceries, the website sections can be called bakery, dairy, and produce (47).
- Online museums, libraries, etc. designed to represent the physical world (1735).
- The Internet Public Library used a metaphor scheme – users could browse shelves, etc. (2403).

Components:

- To succeed metaphors need to be familiar to the users (1168)
- Metaphor schemes can introduce unwanted features (either adhering to the metaphor too closely or breaking from it) (1169).
- You can break the metaphor scheme if you offer services that are not available in the real world version (1170).

Cons:

- These can be taken too far, and end up quickly overwhelming users. It can also get away from designers and compromise usability (2402).

Types:

- Organization metaphors – they recreate the real in the virtual (e.g., an online car dealership will have car sales, repairs & services, etc. sections) (2397).
- Functional metaphors – connects the task you perform in the real with those in the virtual (e.g., in a library website you can "ask a librarian") (2399)
- Visual metaphors – leverage familiar graphics elements and colours to connect the real to the virtual (e.g., the yellow pages online has a yellow background and phone images) (2401).

Hybrid organization schemes:

This when there is a combination of organization schemes used in websites. Generally there is a primary scheme (often topical) and secondary schemes (task, alphabetical, etc.). This allows users to search content on various criteria (e.g., date and location). Large websites typically require several types of structure (686, 1013, 1293).

Examples:

- TV guides can be organized by type of show, subject, alphabetically and time (634).
- Tourism websites – organized by geographical and topical schemes (656).
- Library catalogues use three (on average) organization schemes – you can search by author, title, and subject (1130).

Cons:

- Too many organization schemes mean it's harder to understand the website (1179).
- Hybrid schemes should only be used in shallow schemes (large websites get way too confusing) (1183).

- In large websites if you have hybrid schemes they should be presented separately on a page (preserving the integrity of the schemes) (1185, 1186).

Organization structures:

Organization structures define the relationships between content items and groups. This structure defines the primary navigation systems (1099, 1102).

Hierarchy structure:

This structure is organized with parent-child relationships, aggregating upwards in to broader groupings or going downwards in to narrower groupings. These items can only belong in one grouping (or else it's a polyhierarchical structure, see below). This organization structure (also called a top down approach) allows users to get a handle on the scope of the website (713, 717, 1200, 1202, 2488).

Examples:

- Yahoo!'s curated hierarchical directory (now gone) (968).
- Etsy website, you can find content via category (a small hierarchy) (734).

Components:

- There are two different types of hierarchy structure – shallow and deep.
  - *Shallow hierarchies* mean that you have a lot of high level (or parent) categories, which only contain one level (or child) of categories.
  - *Deep hierarchies* mean that you have a few top level (parent) categories and multiple child categories organized underneath

(352, 714, 1212, 1213, 1214, 1215, 1216, 1223, 2489, 2510)

Pros:

- Particularly good for small websites, they can be used for large websites when the content varies (723).

- Good for websites that have different levels of complexity (the top levels can introduce the topics allowing users to drill down in specificity) (724).

Polyhierarchical structure:

A polyhierarchical structure is organized like the hierarchy structure, but items can be in more than one place. This structure let's designers place things that people expect to find in more than one place and allowing category boundaries to overlap (718, 719, 720, 1203).

Cons:

- If too many items are cross-listed, the hierarchy structure looses its value (1206).

Hypertext structure:

In the hypertext structure, content (any format) is joined together according to the relationships between them by using hypertext links. This is particularly useful if you're adding more content over time. This is a non-linear way of structuring information (745, 751, 752, 1256, 1257, 2502).

Examples:

- The most well known examples of hypertext structure are wikis, and Wikipedia is the best example of those (746, 747, 748).

Pros and cons:

- Success of hypertext structures depend on if the users follow the connections (754).
- The context authors create connections between information, and if they don't know what's related, those connections won't be made (757).
- Organization structure provides great flexibility (1260).
- It's easy for users to get lost in a hypertext structure (following links) (1261).

Database oriented structure:

The database-oriented structure organizes data within a set of relations or tables. Rows in the tables represent records, and columns represent fields. Data in different tables may be linked through a series of keys, just like an actual database (1236). This is for content that has a consistent structure. The individual pieces of content may have no relationship to one another… but they have the same structure, and are made up of the same pieces (728, 735). This allows users to access the content in more than one way (954).

Examples:

- Etsy – all items have the same pieces that make up each item: title, description tags, material, location, etc. Every item on Etsy has to use the same structure (when using the database oriented structure) (732).
- Database structures work for music, product catalogues, books, articles, etc. (anywhere the content pieces have consistent structure) (735).

Components:

- Bottom up architecture (another term for database structure) is suggested by what content is in the website (1007).
- Most of the heavy-duty databases are built upon this database structure (1235).
- Metadata is the primary key that links information architecture to the design of database schemes. It allows us to apply the structure and power of relational databases to the heterogeneous, unstructured environments of websites and intranets (1240).

Pros:

- You just have to store content once and then use metadata (pieces of the structure) to display information in different ways (954).
- This structure is useful for organizing collections of heterogeneous information (1296).

Social Tagging:

This organization structure leverage tags (user or expert created) to provide access to content (790, 798). The relationships (and organization) between content emerge through the tagging efforts of multiple individuals (1268).

Examples:

- LinkedIn uses social tagging to endorse their professional contacts and this allows users to describe their business contacts (1269).
- Delicious and Flickr let users navigate information using the created tags (981).

Pros:

- This structure works very well for very large collections of diverse content, especially where the content readers will have different ideas what it is about (791).
- Tags can help users explore and find related information (792).

Linear pattern structure (from the navigation miscellaneous category):

These aren't that common on the web, a linear pattern is as the name suggests, and one thing follows another in a straight line. These are useful where users need to understand something before moving on (758, 759). This is also called heterarchical (2491).

Example:

- A software instillation wizard (2493).

Hybrid structure:

This structure is made up of two or more of the other structures. Applying multiple organization systems to the same context can allow websites to escape the limitations of using just one structure. Most websites use more than one organization structure (1083, 2505).

Examples:

- A typical website has a hierarchy navigation system, sequencing for signing up, and hypertext links to related content (2506).
- A typical grocery store has a hierarchical aisle system, a heterarchical database to retrieve product information (barcodes), and sequencing for checking out. (2507)

Types:

Hierarchy/Database oriented structures:

This is one of the most common hybrid structures. It's suitable for every size of website. It lets you create a hierarchical sections of the website for basic content and then uses the power of a database to assemble detailed information within a section (761, 762). These structures can be combined in just one section throughout the whole site (or in between) (763).

Pros and cons:

- Database structures can manage large volumes of content, while the hierarchy makes is easier to access (from the top) (762, 765).
- Hard to decide what pieces you'll turn into structures content (the database) and what you'll leave in a hierarchy (764).

Catalogue pattern structure:

This is a database/hierarchy pattern, but more emphasis is put on the database. At the bottom level is the content with up to three levels of hierarchy above (depending on the size of the website and what content you have) (766, 767).

Examples:

- Design gallery pages (955)

Hub and spoke pattern structure:

This is a hierarchy, but used differently than a normal hierarchy. With hub and spoke, people move down one level into something more detailed, return to the starting point (the hub) and then may move to another detail page and then back to the hub (and so on) (768, 770).

Subsites structure:

Subsites (or portals) is a pattern of organizing many subsites in to one website, held together by a homepage or top-level pages. This can use any organization structure in each subsite. These use consistency in navigation, layout, and design to make it seem like one cohesive site (772, 773).

Examples:

- Universities often use subsites because they have a variety of content and many different audiences (775).
- Subsites can be used for government websites (776).
- eBay motors has subsites with highly specialized navigation systems (991).

Focused entry points:

In this hybrid structure a series of entry points are provided to help users find their way (around a hierarchy, usually). They don't have to cover all the sites content, just key information (777, 778).

Labeling Systems:

Labels are the things that define chunks of content. The goal of labels is to communicate information without taking too much space of the website. Labels should be designed for the user and they should be clear. Labeling shows the organization scheme of a site (794, 1317, 1324, 1512). Although, one needs to remember, labeling is intensely impacted by the choices their authors make (1094).

Components:

- Often successful labels are invisible, they don't get in the way of the user (1325).
- Your labels will never be perfect; there will always be people who misinterpret them (labels are not one size fit all) (1467).

Types of labels:

The types of labels used define and infer the content they represent (1103).

Textual labels:

Textual labels are the most common type of label, and can be used to convey messages and brands of a website (997, 1348, 1514). However, when we use words as labels for our categories, we run the risk that users will miss the meaning (1084). An example of textual labels is "Contact Us" a label that represents chunks of text, which we know what is represents (1315).

Contextual Links:

Contextual links connect chunks of information to other information on different pages (or on the same page) (1349). These appear as hyperlinks within the text or content of a page (1358). These links are easy to create, the content authors generally develop them, but information architects can offer guidelines to how this can and should be done (1358, 1359, 1361, 1378). Contextual links are given context by its surroundings (in the text); this does rely on a good content author though (1365).

Example:

- GOV.UK contextual links draw on surrounding text for context (1366).

Cons:

- Contextual labels are personal; we may think we will be taken to one place but end up in another (e.g., the contextual label "Shakespeare" doesn't take you to the bards Wikipedia page rather to information about a town in Arizona) (1364).

- Contextual links can be ambiguous depending on the audience (a blog has more leeway with vague contextual labels) (1368).

Headings:

Headings are labels that describe chunks of information and help keep webpages orderly. They also help the user from feeling overwhelmed by the choices offered (75, 797). They are just like print headings, in that they describe the content that follows them (1350). Heading labels are often used to establish the hierarchy within a webpage (1379).

Example:

- In hybrid structures, headings help us to distinguish a websites subsites (just like in a book were headings help to distinguish chapters) (1381).

Component:

- Headings are very important in a hierarchy. Hierarchy relationships between headings are established visually and these headings are given meaning when they are found within a hierarchy (1383, 1385).

Navigation Labels:

They represent the available options in the navigation systems (1351). Navigation labels show the organization and navigation schemes of your site (1318). Navigation labels help provide context to a website (you see certain labels you know where you are, e.g., navigation links called "Loans and Credit, "Investing," and "Wealth Management" we think bank) (989).

Example:

- When developing the "Boxes and arrows" website the navigation labels were shifted from types of content (How-tos and interviews) to disciplines (Information architecture and interaction design) (571).

Pros and cons:

- Navigation labels are used throughout the website, so if there are problems with the navigation labels they will be experienced repeatedly (1406).

Index terms:

Index term labels (also called keywords, tags, descriptive metadata, etc.) can be used to describe any type of content: sites, subsites, pages, content chunks, etc. and also represent content for searching and browsing (1352, 1423). Using index terms created from controlled vocabulary or thesauri has more value (1446).

Examples:

- The index of the SFGate website is generated using index term labels, which in turn are used to identify content from many different sections of the site (1438).
- Searching for the embedded index terms would return results even if they were not in the page's text (1443).

Icon labels:

Icons labels are labels that are not represented by text rather they are images/graphics. Icons are typically used for navigation systems and organization systems, where the list of options is small (1458).

Examples:

- Shopping carts in commerce websites that are represented by a small shopping cart image are an example of icon labels (354, 1329, 1996).
- A business/organization logo is an example of icon labels (1066).

Pros and cons:

- Icon labels are more limited in expression than text (1457).

- Icon labels add aesthetic appeal, but you need to ensure that they don't hurt usability (1463).
- Icon labels are especially useful as representational shorthand labels (1465).
- Unless you have a patient and loyal audience who are willing to spend time learning icon labels, you should limit your use of them (hard to immediately understand them) (1466).

Labeling consistency:

A good label is so obvious it's dull as dirt. A good label doesn't make you pause (and it never makes you think) (97). Consistency brings all the labels together and it means that the site is predictable, if you see one or two labels that are the same then you know what to expect (1476, 1477). Importantly the labels need to be consistently applied throughout the website (this even applies to their colours and locations) in order to build a sense of familiarity with the user. For example, using the label "Main" on one page, "Main Page" on another, and "Home" elsewhere is a lack of consistency site wide and could destroy the familiarity that the user needs (1408, 1409, 1411). Confusing and inconsistent labels can negate all the investments made in the design of the website (1344).

Examples:

- A school website has a section called "Parents", which may contain specific news for parents or ways for parents to be involved in the school, but it actually contains cafeteria and after-school care information. The label doesn't communicate what is included in the section (822).
- Starbucks uses inconsistent labels that don't represent the content they link to. For example, what is the difference between "Coffee," "Coffeehouse," and "Shop." (1339)
- In budget car rental, they have inconsistent labels for the contact page – one is labeled "Contact Us," and on another "Customer Care." (1490)

Components:

- Terminology: Use the most correct terminology you can use, balanced with what the audience knows. Correct terminology will describe the content well and "educate" users on what the correct terms are (802, 803).
- Be careful using jargon for you labels and if you do use it, only use jargon that your audience knows (814, 815). This also includes organizational (business) jargon, which will only work for the .01% of users that work at the business (1342).
- You need to plan for labels to change over time (as your audience and content evolves). Languages also change over time, so consistency needs to be monitored as these are updated (830).
- Style - Haphazard usage of punctuation and case is a common problem within labeling systems, and can be addressed, if not eliminated, by using style guides. (1479)
- Presentation - consistent application of fonts, font sizes, colors, whitespace, and grouping can help visually reinforce the systematic nature of a group of labels. (1480)
- Syntax - It's not uncommon to find verb-based labels (e.g., "Grooming Your Dog"), noun-based labels (e.g., "Diets for Dogs"), and question-based labels (e.g., "How Do You Paper Train Your Dog?") all mixed together. Within a specific labeling system, consider choosing a single syntactical approach and sticking with it. (1481)
- Granularity - Within a labeling system, it can be helpful to present labels that are roughly equal in their specificity. Exceptions (such as site indexes) aside, it's confusing to encounter a set of labels that cover differing levels of granularity—for example, "Chinese restaurants," "Restaurants," "Taquerias," "Fast Food Franchises," "Burger Kings." (1482)
- Comprehensiveness - People can be tripped up by noticeable gaps in a labeling system. For example, if a clothing retailer's website lists "trousers," "ties," and "shoes," while somehow omitting "shirts," we may feel like something's wrong. (1484)
- Audience - Mixing terms like "lymphoma" and "tummy ache" in a single labeling system can also throw people off, even if only temporarily. Consider the languages of your environment's major audiences. If each audience uses a very different terminology, you may have to develop a separate labeling system for each audience, even if these systems are describing exactly the same content. (1486)

Navigation Systems:

Navigation systems chart the course, determine your position and find out way back; they provide a sense of context and comfort as you explore websites (1747). They seem simple, but they are the most subtle and complex part of the interface. The job of navigation is to clearly state where a user will travel in the information architecture (364). Navigation systems should help users do what they want, but also it needs to help them do what you want them to do (500).

Global Navigation:

Global navigation is a set of navigation tools that are consistent throughout the site; they allow users to find their way easily through the site by providing these constant links (6, 7, 386, 1036, 1570, 1571, 1711). Global navigation tells you what the site thinks you might want and what the site is about (388). This type of navigation is so important to get right because this determines how and what users think of your website and if you get it wrong it has a huge impact on the usability of a site (the problem will be on every page of a website) (391, 1580). Global navigation is almost always located near the top of the website because that allows you to focus the entire rest of the page on content (sometimes it can be found along the left or right side of the page as well) (392). Global navigation also shows the first level of the organization structure for your site (426).

Examples:

- Global navigation of three hotel websites (Omni Hotels, Holiday Inn, and Park Lane Guest Suits) all have the link “Reservations” in their global navigation toolbar because it is the one thing they all want you to find (8, 10).
- Fancast’s global navigation links include, TV, Movies, and People (502).
- Smithsonian [older version] successfully combined topics and tasks on the global navigation horizontal bar (1181).

- Apple and Acer's global navigation reinforces the sites structure and provides contextual clues to identify the users current location. Others like Dell have a much simpler global navigation structure (and doesn't do either) (1574).

Components:

- Horizontal navigation bars show the categories of a site, stretched across the page (usually right at the top or beneath a banner/logo. These are suitable when you have a small number of top-level items that can fit across the screen (859, 860, 861).
- Vertical navigation bars sit at the left or right side of the website. These are most useful when you have more top-level groups than would easily fit across the top of the screen, you want to add more labels over time, if you have long labels, and if your hierarchical site only has a few levels (you can show the hierarchy by indenting subcategories beneath the main categories) (866, 867, 869, 870).
    - Right vertical navigation bars let the content be the focus of the page and can be just as useful as left vertical navigation bars as long as they are easily recognizable (872).
- Combining vertical and horizontal navigation bars you get the inverted "L" navigation, especially useful for larger sites (873, 902).
- Drop-down navigation menus use the horizontal bar and when someone hovers over the main category and menu appears containing the second level categories (876). The advantages of these is that people can see the website structure and let them navigate more efficiently (877, 878).
- Flyout navigation is similar to the drop-down menu, except it's used in the vertical navigation bar and shows three levels of categories (879, 881). Though these can be hard to use depending on the audience of your site (880).
- Mega menus (or large drop down menus) are becoming very common, and allow users to jump deeper into a site without clicking through every level. And if they're grouped and laid out well they show multiple levels of the site (884).

Local navigation:

Local navigation appears when you are within a websites hierarchy and displays sublevels of the pages that are near to where you currently are in the sites (376, 1037). Local navigation can link to the other pages within the same section (411). Local navigation helps in two ways (a) they aid users in seeker tasks, and (b) it helps users browse more specific topics (403, 405, 412, 413). Local navigation is also called section navigation, sub-navigation, and subsites (407, 424, 1592).

Examples:

- When you click on "Research & Strategies" in the Charles Schwab website, seven additional links appear – "Markets," "Stocks," "ETFs," "Bonds & Fixed Incomes," "Market Insight," and "Portfolios." These are the local navigation options for the main category (399, 400).
- The Sapient Interaction website has local navigation within the middle of the page (in the Services section) (410).
- On the Charles Schwab site the section "Mutual Funds" has two local navigation schemes (one nested beneath the other) (425).
- Large sites like GE.com often provide multiple local navigation systems that have little in common with one another (1590).

Cons:

- There are many bad examples on the web, where different section of local navigation varies in style, design, etc. because multiple groups (who control the subsites) have run amok.

Contextual Navigation:

Contextual navigation (also called associative navigation) links together related content, often embedded within text and generally used to connect highly specialized content within a website (454, 907, 1042). Contextual navigation supports associative learning by creating links between information that users can follow (1604). Contextual navigation occurs in the context of the content and explores relationships between items (1754). It addresses what happens once the user has interacted with the website and it stops users from leaving when/if they don't find what they want (441, 442). One important consideration when dealing with contextual links are to make sure that they are well describes and the destination is clear (894).

Examples:

- "See Also" and/or "Related Links" are common forms of contextual navigation (generally located at the bottom of a page)(65, 66, 179, 576, 895, 896, 898, 899).
- The Huffington post puts related links at the bottom of articles (e.g., "More in Politics…") (439).
- Comcast has contextual links in the form of "Most Viewed," "Recommended," and "Emailed" stories (443).
- YouTube offers the options of viewing similar videos once you have finished watching one (444).
- The website "Boxes and Arrows" added contextual links to the bottom of the articles published on their site, preventing "dead ends." (579).
- Wikis are one type of site that uses contextual navigation (relies on them for main navigation system) (891).

Pros and cons:

- Contextual links are the most important types of links because they drive most of the usage of the site (440).
- Contextual links stop a page from being a dead end (they provide additional options for the user) (445).
- You need to make sure that contextual navigation is visible and distinct from the rest of the content (but still reflects your website design) (892, 893).
- Generally content authors manually add the links to the content (subjective decisions) (1606).
- Inline contextual links may not work because users generally scan the information too quickly, so you may want to design a specific area in the webpages for the contextual links (1610, 1611).
- Moderation of contextual links is key – too many can clutter up a webpage (1613).

- You can automate the creation of contextual links if you have a very large site (using metadata) (461, 2424)

Types:

- Link items by time (when they were published, when they occur, etc.). On news websites there could be earlier or newer stories related to what you are viewing (455).
- Link articles by type (more article, more videos, more photos, etc.). YouTube links videos together (456).
- Link items by subject (more items in the same category). Huffington post links to more articles based on the subject of the article being viewed (457).
- Link items by interest (most popular items). Comcast does this offers links to most viewed (458).
- Link items by owner or group (link articles around the author) (459).
- Link items by community (items linked together based on what you liked or what people like you liked) (460).

Breadcrumb Navigation:

Breadcrumb navigation is a visual indicator that tells users where they are in a website (3). Generally located near the top of a page (below the global navigation but above the content), they show the users where they are in the hierarchy of the website (64, 908, 910, 914). You can use the breadcrumb navigation to move up and down the hierarchy of the site, by clicking on one of the categories (see example below) (432, 911).

Example:

- On Gap's website, they highlight what section Austin's in (Men > Accessories > Scarves) (64).

Hypertext navigation:

Very similar to contextual navigation (more of a subsection of it), hypertext navigation transports users into the middle of an unfamiliar site (e.g., by passing the homepage) (1557). It also lets users move anywhere within a site (up down, left, or right in the hierarchy), which they want to be able to do (move both laterally and vertically in a site) (1563, 1564). This provides flexibility to a website (1697).

Components:

- Make hypertext navigation links change if the users have already clicked on them (15).
- If it is a link, make it look like a link (don't get fancy) (61).
- If there is two hyperlinks on the same page that links to the same content, make the labels the same (or at least very similar) (808).

Utility navigation (from navigation miscellaneous category):

Utility navigation connects pages and features that help visitors use the site itself. This includes features like sign-in and access to user information (profiles or credit card information). This is all the stuff that lies outside the main content organization, yet is critical to the site's functioning (378).

The best way to describe utility navigation is detailing what it includes: sign-up, sign-in, access a user's account or profile, help, contact information, links to physical locations and can also include jobs, blogs, press releases, bookmarks, favorites, and history (470, 479). Even though these links are important and necessary, they're usually the last to be considered, though they serve as lifeboats for visitors who arrive at your site and find themselves overwhelmed with information (471). Utility navigation handles the items that are not covered by global, local, and contextual navigation (483).

Examples:

- On the Charles Schwab site the utility navigation includes "Log In," "Contact Us," "Visit Us," and "Search" (467).

- E*TRADE utility navigation includes “Log On,” “Open an Account,” “Customer Service,” and “Search” (478).

Supplemental navigation:

Supplemental navigation provides different, but complimentary ways of finding information and completing tasks (compared to the other navigation methods mentioned above) (1544, 1642). This is the “emergency backup” of a website to support search and browsing systems (1650).

Sitemaps:

Sitemaps are a condensed overview (a single page that listed all pages) and links to major content areas within the environment, usually in outline form (918, 1038, 1545). Sitemaps used to be quite common on the web, but disappeared for a while (they are reappearing again) (919). Sitemaps have two purposes (a) they work for humans who want to see all of a website with one glance, and (b) search engines spider sitemaps to return results (919, 920, 1663). How to format sitemaps differ depending on the website, some say sitemaps should list as many pages as possible (912), but for large sites you have to decide what to list (922). A typical sitemap provides a broad view of the content in the system and allows people to access random sections of a site (1652).

Pros and cons:

- Make sure that you have a way of maintaining the sitemap (they can be hard to maintain), especially when working with dynamic content (923, 1773).
- You should consider the system’s size when deciding to use a sitemap (you don’t really need one for a small site) (1659).
- The design of a sitemap affects the usability of a website, the best way to create one is to reinforce the hierarchy of the website (1660, 1661).

Indexes:

Indexes (organized alphabetically) are one of the best ways to find things on a website, especially for known-item finding (they let users jump straight to the content) (589, 930, 949, 1667). They provide a list of links to the contents of the website (1039, 1665). These also help people understand the difference between their terminology and technical terms (931). Indexes are a good back up, if you can develop or don't have a good search system (1763).

Examples:

- SFGate website index is generated from index terms labels, which are used to identify content from many different sections of the site (1437).
- The United Nations presents a comprehensive alphabetical index (1668).
- Comcast XFINITY website presents a simple site index (1675).
- The Center of Disease Control and Prevention two-step site index features term rotation and see/see also references (1681).
- The Michigan State University site index has hundreds of the sites best bets results and renders them in an alphabetical list (1682).

Components:

- You should list terms (that have two common terms) under two headings, but be conservative (add only when you absolutely need to) (646).
- You don't usually need to list every content page in the website, you will probably want to include all the main topical pages (depending on how the site is structured) (647).
- Good indexes match the terms that their users think of (or at least provide two common terms, see above) (932).
- Good indexes also provide good coverage of the content (maybe not every page, but every topic) (933).
- You may want to index paragraphs, it depends on website (1678).
- You can use controlled vocabularies to create indexes (automatic generation using indexing), but these should have drop-down menus for users to choose the correct term (1684, 1687).
- If you have a small website you can create an index manually (but be careful!) (1680).

- Term rotation (or permutation) is a useful tool in indexes; this rotates the words in a phrase so that users can find the phrase in two places in the alphabetical sequence (e.g., “Abuse, Elder” and “Elder Maltreatment” are used in the CDC index) (1688).

Guides:

Guides are supplemental navigation systems that provide specialized information on specific topics, as well as links to related subsets of content (1040). Guides can take several forms: step-by-step tutorials, guided tours, walkthroughs focused around a specific audience, topics, etc. (976, 1690). Guides are useful for introducing new users to the content and functionality of a website (1692). This includes using it as a marketing tool – introducing users to the restricted portion of the website (1693). Often the format of guides includes screenshots combined with text that explains the steps (section of the website) that are being described (1698). These are generally not used that much (1707).

Example:

- Guided navigation is embraced in online retail (clear link between flexibility and profitability) (2345).

Components:

- Guides should be short (1700).
- Users should be able to exit the guide at any point (1701)
- Guide navigation should be consistent (previous, next, etc.) (1702).
- Guides should be designed to answer questions (1703).
- Screenshots should be clear, crisp, and enlarged to show details (1704).
- If it has more than a few pages, the guide should have a table of contents (a text version not sitemap) (1705).

Configurators/Wizards:

Configurators and wizards are a specific supplemental navigation tool, which lead users through sequential sets of steps (and may also link to related subsets of content) (1041, 1708). Wizards are used to accomplish a goal that has many steps; they are linear so they make sure you don't miss a step (configurators let users jump around though) (330, 331). They are useful when you don't have to perform a task too often (they can be plodding and slow) (332).

Examples:

- Amazon's e-commerce checkout is an example of a configurator/wizard. Users need to confirm their order, enter payment information, and tell you where it's to be shipped – each are on a separate page that cannot be skipped (514).
- Sophisticated configurators, like Motorola's Moto Maker allow the user to easily traverse complicated decision-making processes (they can move back and forth, jump between steps, etc.) (1709, 1710).
- Apple makes the changes to the image of the product that you are configuring (1713).

Components:

- Wizards are a good choice when the audience is not tech savvy or when the Internet connection users have are slow (333, 334).
- You need to provide the users clear options (they don't always know what options exist within a configurator/wizard)(1712).

Control Panels (from navigation miscellaneous category):

Control panels are forms all in one page (with complex layouts) (337, 338). It's a good choice to use these when the audience is tech savvy with fast internet, when the application is easy to understand, when you need to give context to the object and when they are used a lot (339, 341, 342, 343).

Toolbars (from navigation miscellaneous category):

Toolbars keep the tools for interaction close to the workspace they affect. This is useful when you are frequently tweaking something (writing or drawing) (345). When you find that there are several links joined together you can form them into a module (Flickr has picture toolbar modules) (360).

Example:

- The Flickr toolbar is made up of items that will help you edit a picture so grouped together into a module (361).

Components:

- Use toolbars when there are many steps in a task and the steps can be done in any order (346).
- Use toolbars when things need to be undone and redone and well as just plain done (347).
- Use toolbars when the proximity of tools to the workspace is important to the task (348).

Pagination navigation (from navigation miscellaneous category):

Pagination is a special form of navigation; it's a simple tool that lets people flip through multiple pages (breaking a large group of items into bite-size pieces) (503). This prevents information overload (504). This should be taken away when the user goes to print an article (507). Additionally, there should be a "View All," option (for smaller articles/pages of information) so that all the information can be viewed on one page if the user wants (508, 509).

Examples:

- Amazon has a nice pagination design. It tells you what page you're on, offers links to the previous and next page as well as links to specific pages (505).
- Fancast provides users with the option to view more or view less results per page (506).
- Boxes and Arrows splits its articles across several pages, however, the number of clicks per page decreases each page because only half of users will continue to read (510).

Advanced navigation:

Good navigation design should be tackled before you move on to advanced options (and these should not be tackled by novices).

Personalization:

Personalization provides information to users based on their behaviour, needs, etc. (1720). In short, with personalization, websites guess what the users want (1722). Personalization also extends to searching, a website can use your social, geographic and demographic context to personalize your search results (309, 312).

Examples:

- Facebook learns more about you overtime to personalize your news feed (533).
- Amazon.com uses personalization (remembers names, addresses, cards, etc.) (1727).
- Netflix recommends users based on what you have watched as well as what other similar users have watched (by comparing ratings you have given to movies/TV shows) (313).

Pros and cons:

- Personalization starts to break down when they start to recommend items (don't know what you've purchased elsewhere, just in that site) (1727).
- Often users don't have time to teach the system, or because we want to maintain privacy – this will hinder personalization (1728).
- Personalization works really well in limited contexts, but you shouldn't overreach (e.g., use it to drive the entire user experience on your site) (1729).

Customization:

Customization gives the user control over their section of the site, users tell the website what they want (1721, 1723). Users should be able to customize their website profiles at will, as their lives change (replacing photos, etc.) (530).

Examples:

- MySpace allows users to customize their pages (down to the HTML) (531).
- LinkedIn allows users to customize their relationships between other members (535).
- Orkut allows users to customize relationship connections between users (with terms like "friends," good friends," etc.) (536).
- Facebook wants users to define their relationships with your "friends" to provide you with a more customized experience in their site (new stories, etc. that you have a greater interest in).

Pros and cons:

- Customization has both promises and perils – giving users control alleviates pressure on design (1730).
- Customization delivers value, for example Gmail lets you customize the look of your email (1731).
- Problems with customization are that users don't really want to devote that much time, unless it's important to them (1732).
- Customization works for corporate intranets (they have a captive audience who repeatedly use the site) (1733).
- Customization works great for tracking sports, but not so much for broader news and research needs (1734).

Visualization:

Visualization uses images to navigate, this is especially useful when users must select among a result set of elements by their looks (images), and in the case of shopping for physical goods (1737).

Examples:

- Maps used to display information are a form of visualization (as long as you can click on the map to navigate to content) (658).

- Tag clouds are another type of visualization that represents the tags used in your site (the larger the word the more times it appears in your site) (916). These can be misinterpreted and confusing to some users (917).

Social navigation:

With the rise of social media, social navigation has become an important approach for finding information based on user interest (1738). Tags allow users to move fluidly between objects, authors, tags and indexers. And when many people get involved with tagging, opportunities arise to transform user behaviour and tagging patterns into new organization and navigation systems (1267). This is built on the value of users observing other users (especially users that you have some sort of relationship with) (1739). Social navigation can help users discover content based on its popularity (volume of traffic or some kind of voting system) (1740). The popularity of this navigation method depends on the popularity of social media (1744).

Examples:

- Delicious aggregates the most popular websites and tags across all their users, making their front page a guide to the newest cool stuff on a variety of topics (216).
- Digg an online news site, which promotes stories to the front page using reader, votes (529).
- Reddit, a content aggregation and discovery service, employs a voting system for front-page content (1741).

Search Systems:

Search systems allow users to search various types of information environments (from the entire Web to a small website) (1759). Part of the value of search is that they can be comprehensive and can cover huge amounts of content quickly (1796). Search systems help when a website has too much information to browse, if the site is fragmented, it is a learning tool, it should be in a website because users expect it to be there, and search tames high dynamism in websites (1765, 1766, 1770, 1771, 1772).

Search algorithms:

Search algorithms determine which content matches a users query (Google PageRank is one famous example), it processes queries into something the search software understands (1052, 1781). They balance fast, easy, and magic to create a search algorithm for your site (designers have to choose what to give more emphasis to) (265). There are tons of search algorithms, but they all work towards identifying the best pool of documents to be returned (1858, 1893).

Search algorithms can influence the results of a search in many different ways, for example you can weigh search results by the prices attached to them (31).

Examples:

- Pandora uses the Music Genome Project to power their recommendation engine (37).
- Most search algorithms employ pattern matching (1860).
- Cite Seer's algorithm finds related citations for an article you like (1892).
- PageRank (part of Google's search algorithm) is a famous example (1967).

Recall and Precision:

Recall is how good the search system is at finding absolutely everything you were searching for, and precision is how good it is at organizing these results by how relevant they are to your query (243). People looking for a particular answer prefer high precision, it doesn't matter what's returned as long as it includes the right answer (1871). Recall and precision is inversely related (meaning that you can't have both high recall and high precision in one search engine (1872). Choosing between recall and precision depends on your user (1883).

Examples:

- As an example, users who are "ego-surfing" will want to see every mention of their names – they're hoping for high recall (1869).

- Searching for “William Faulkner” in the author field will result in higher precision, assuming we’re looking for books authored by Faulkner” (1884).

Pros and cons:

- As the catalogue in your website grows, it is harder to get precision right (247).
- In a smaller catalogue precision is easier to deal with (248).
- A problem with recall is that along with the good comes the irrelevant (1870).
- High precision and high recall is the ideal, but you have to trade off one for the other (2122).

Search interface:

The search interface is where users enter and revise their search query. There is so much variation among users and search technology functions; there can be no ideal search interface (1043, 2001).

Search interface components:

The two main components of the search interface are the search box and the search button (2014). It’s best to keep you search interface as simple as possible, for example don’t present it along other boxes and have just the one search box (2014, 2024, 2025). Consistent placement of the search box and button site wide is best practice for information architecture design (2026).

Advanced search:

Advanced search interface provides more manipulation of the search system and are typically used by two types of users: advanced searchers (doctors, lawyers, librarians, etc.) and frustrated searchers (advanced searchers are the norm) (2036, 2038). Advanced search provides flexibility and power to expert users (2039). If you need an advanced search add one, but the goal is that users never need to use it (2041).

Example:

- The Advanced Wine Search provides the ability to combine facets into the rich type of query (2333).

Search results:

There are many different ways of presenting search results, information architects need to decide what the users want (and what fits with your content) (1903, 2071). One option is to present results that have been indexed with the same metadata (for example the DuckDuckGo search engine offers more matches for the search terms in the same domain as that result) (1889).

Displaying results:

This is the presentation of content that matches the user's search query (1054).

Format:

Formatting search results includes making decisions about how many documents to display on one page, making the links look clickable, etc. (276, 1923). For example, you should err on the side of simplicity when listing the returned results (showing a small number) because you do not know the circumstances of the users computer, Internet connection, etc. when they are on your site (1926). Also consider that if the results your users are looking for aren't on the first page, it might as well not exist (they rarely go past the first page) (511).

Examples:

- Google displays results in a triangle patter (from the top left corner), allowing users to quickly select a result worth clicking on (274).
- Yelp iPad app allows users to decide how they want to view the results (map or list) (1912).

Information displayed for retrieved results:

You can display different types of information about the retrieved search results (e.g., page title, extract, and date) (1019). What you decide to display depends on your users, users who know what they want, would like representational content (title, author, date, etc.). Users who don't know what they want would benefit from descriptive content (abstracts, keywords, etc.) (1908, 1909, 1910).

Examples:

- In phone directories it makes sense to show the phone number in the results because that is what users expect to see (1919)
- The Verge highlights the search terms in the retrieved results (1922).

Components:

- Highlight the search query within the information returned; searchers often look for this to determine what result to choose (277).
- Snippets of information about the retrieved results come from the content of the items themselves (this helps the users decide if it is a good choice) (279).
- When it's hard to distinguish between results show more information to help differentiate (1913).
- If you don't have a large amount of content you can display more information if you think it would be helpful (1916).
- What you display depends on the type of content information that you have to display (e.g., how your content is structured) and on how the content will be used (1918).

Sorting results:

Sorting is the act of arranging search results according to rules (2496). Sorting search results is especially useful for users who are looking to make a decision or take action. For example, users who are comparing a list of products might want to sort by price or another feature to help them make their choice (1941). Any content component can be used to sort, but it's sensible to provide users with the options to sort on components that will actually help them accomplish tasks (1942).

Alphabetically sorting:

Just about any content can be sorted alphabetical (by any number of content components – title, author, department, etc.) (1937, 1946). Everyone understands the alphabet so this is an easy way to sort results (1947). Though you should remove the initial articles (a, the, etc.) when sorting them alphabetically (1948).

Example:

- Users are likely to look for “The Naked Bungee Jumping Guide” under “N” rather than “T” (1948).

Chronological sorting:

You can sort your search results chronologically (by date), this is useful for content that is time sensitive, press releases or other news oriented information, and historical data (1936, 1949, 1951, 1952). You can sort chronologically in reverse order as well (1951). When sorting by date you can also draw on the built-in dating information of the content (1950).

Ranking results:

Ranking search results is helpful when there is a need to understand information or learn something (1943). Deciding on how search results are ranked depends on your users information needs, what results they are looking for, and how they would like to use the results (1933).

Popularity ranking:

Ranking results by popularity can be done in a number of ways. Google ranks articles by popularity by measuring the number of links that link to it (1966). Large sites with lots of content can better take advantage of popularity ranking than smaller sites (1968). Smaller systems don't have enough variation in the popularity of different documents (1969).

Example:

- Google ranks results by popularity, it does so by factoring how many links there are to a retrieved document (1965).

Users' or experts' ratings:

If users are willing to rate the content on your site, it can be used to rank search results (1970). Most sites don't have a sufficient number of motivated users to employ valuable user ratings. However, if you have the opportunity to use this data, it can be helpful to display user ratings with documents (1973).

Examples:

- Yelp! uses the expert/user ranking system, this is integral to helping users judge the value of an item (1971). This works for Yelp! because it has users who are willing to rate things (1972).
- Wine.com has added ratings from several magazines (WS= Wine Spectator, etc.) that let users sort results by (2339).

Pay-for placement:

Advertising has become the predominant business model for publishing online, so it is no surprise that pay-for-placement (PFP) has become a commonplace in many search systems (1975). If you have lots of advertisements on your website you should consider pay-for-placement. Users may think that those who can pay are more stable then the other options in the search results (1978).

Example:

- The first result in Yelp! Is actually a paid for advertisement (1977).

Relevance (from the search result rankings miscellaneous category):

Generally ranking search results by relevance is the default (from most to least) (1944). Ranking by relevance is relative, and be careful because users always assume that the first results are the best (1945). Different relevance ranking approaches make sense for different types of content (1959).

Examples:

- Document A might be ranked higher than Document B, but Document B is definitely more relevant. Why? Because while Document B is a bibliographic citation (for a really relevant work), and Document A is a long document that just happens to contain many instances of the terms in the search query (1960).

Components:

- Relevance ranking can be determined by the presence of a term (1954).
- How frequently those terms occur in that document (1955).
- How close together those terms occur (e.g., are they adjacent, in the same sentence, or in the same paragraph?) (1956).
- Where the terms occur (e.g., a document with the query term in its titles may be more relevant than one with the query term in its body) (1957).

Best Bets:

Best bets are when human beings go in and muck with the search results, adding in new results by hand (300). Users might assume that these search results are automatically generated, but humans are manually modifying the information architecture in the background (1022). Because best bets are added manually, they should only be done for the most popular searches on your site (save time or money) (1964).

Examples:

- HP has weighted the search results for "Digital Camera" to include best bets (301).
- REI uses best bets for promotion (You're looking for GPS? We got them on sale) (302).
- The "Editors Choice" results in BBC are manually created and assigned to search terms (1021).
- Michigan State uses best bets to populate their index (1683).

Hybrid search results (clustering):

Mainly you use existing metadata to cluster search results, like document type (.pdf, .doc, etc.), topic, audience, language, etc. to divide the search results into clusters (though these can be costly) (1985, 1987). Clustering provides content for search results and lets users select the best category for their search (1990).

Example:

- Forester uses a hybrid approach to clustering search results by contextualizing the query "user experience" with roles such as "Marketing Leadership" and specific date ranges (1989).

Additional actions:

Additional actions come into play once a search has been conducted. The best approach is a "no dead end" policy to address problems, this means that users always have an additional options (even if they've retrieved no results) (2061). These options include search tips, human contact number if it doesn't work, etc. (2063, 2066). If the results of the search are not satisfactory you can also explain what happen behind the scenes, providing the user with a better understanding of what happened (2049).

Example:

- The New York Times site provides an excellent example of explaining to the user what just happened in the search (2057).

Save search:

In some cases, it's the search itself, not the results that you want to save (1998). Saved searches are good in dynamic domains and when you want to manually re-execute a search (1999).

Example:

- Save search options are generally located in the upper right corner of the returned results (2000).

Select a subset of results:

Sometimes when you are searching you want to "shop" through the results – saving a subset of results (1994). If you're sorting through dozens or hundreds of results, you may need a way to mark the documents you like so you don't forget or lose track of them (1995).

Example:

- Users can save search results and come back to them ("browsing" the shelves) (1997).

Narrowing results down:

For large search results you can provide alternative search terms to narrow down the results (so they don't go to the competitors) (453). These can include date filters (2053). Or allow users to search within the search results (2060).

Example:

- After the initial search for hotels in New York City retrieved over 600 results, we can then filter by hotel name for particular brands to narrow our retrieval (2060).

Repeating/new search:

Provide the option of revising users search results, especially when the search results are too large (1930, 2059). You should provide instructions on how to revise your search, or keep the search term in the search box when results are returned so that they can change it (2019, 1931, 2046).

Example:

- Reuters keeps the search query in the search box so users can search again. Letting users modify their search without re-entering it is helpful (1931, 2043).

Query builders:

These are ways of enhancing a query's search performance; they are often invisible to users (who may not value or how to use them) (1047, 1895). Query builders have pros and cons, which address different information needs in different situations (1902).

Spell checkers:

Spell checkers allow users to misspell terms and still receive results by automatically correcting the search term (1896). Almost everyone misspells, so this feature should definitely be considered for your search engine (1901). Google and Yahoo! created their spell checkers by monitoring when a query has no results and then looking at the following search they conducted (294).

Examples:

- According to Zabar's website they don't have "chedder," except they do only it's called "cheddar." (172).
- Yahoo! Recognizes the wide variety of spellings humans come up with (173).
- We type "fuschia," but we don't see anything on the page so we don't click on anything then redo the search with the correct spelling (295).

Phonetic tools:

Phonetic tools are useful when searching for a name. They can expand a query on "Smith" to include results with the term "Smyth." The best-known example is Soundex (1897).

Stemming:

The search might provide automatic stemming, which expands a term to include other terms that share the same root (or stem) (1875). Strong stemming expands the search query, no stemming means that variants are ignored, and weak stemming might just pluralize the term (1877, 1879, 1880).

Example:

- If the stemming mechanism is very strong, it might treat the search term "computer" as sharing the same root ("comput") as "computers," "computation," "computational," and "computing" (1876).

Natural language processing tools:

Natural language processing tools examine the syntactic nature of a query – for example, is it a "how to" question or a "who it" question? – And use that knowledge to narrow retrieval (1899). These components "feed" other components, such as a thesaurus that's used to enhance search queries (1069).

Example:

- Siri uses natural language processing to figure out if it should trigger a web search or a bad joke (1899).

Autocomplete/Autosuggest:

Autocomplete/Autosuggest works as the user types in their query, they populate a drop-down menu with options based on what you've already typed. If your website is large and gets a lot of traffic and repeated searches, autosuggest would work well (285). These are widely used patterns for interacting with search systems (2028). Displays for autocomplete/autosuggest range from simple (straightforward text lists) to complex (popovers with highly customized layouts) (2031). These are very useful for users who only have partial information (2032).

Example:

- The reigning query helper is Yahoo! Search Assist (it finishes users queries and provides related concepts to explore) (288).

Query Language:

How users compose a search query varies widely, though Yahoo! And Google showed that 80% of users searched using one or two-words (126). Early search systems involved very complicated query languages (see the components below) and these still work, but improving search systems mean that these aren't necessary and the average user wont use them (2006, 2013).

Components:

- Boolean operators (AND, OR, and NOT) are one type of tool used during users search queries (though mainly by "experts" – librarians, researchers, etc.) (2006, 2013).
- Search engines can strip out "stop words" (like a, the, and of) to help a users search query (1020, 1888).
- Proximity operators (e.g., ADJACENT or NEAR) (1046).
- Or ways of specifying which fields to search (e.g., AUTHOR="Shakespeare")(1046).

Indexing content for searching:

Search engines index the contents of a site and information associated with each document (like author, titles, controlled vocabulary terms, etc.), but it can also exclude sections (like an articles citations) (1775, 1778, 1798). It is a good idea to index your content because the search results will be better and you can choose to exclude administrative data, etc. (1849). Choosing what to index in your information environment is an important step when configuring your search system (2069).

Example:

- Yelp! Indexes many different components of the items for searching (business name, ratings, operating hours, links, etc.) and leaves out items like user reviews (1850, 1851)

Search Zones:

Search zones are subsets of site content that have been separately indexed to support narrowing searching (e.g., searching the tech support area within a software vendors site) (1053). When users search a search zone, he has, through interaction with the environment, already identified himself as interested in that particular information (1803). Search zones should correspond with your users needs and retrieve more relevant results (1804).

Example:

- Windows 8.1 lets users select search zones based on the type of content they are looking for (Settings, Files, etc.) and by its location (Web images, Web videos) (1805).

Pros and cons:

- You can create search zones in as many ways as you can separate documents (1809).
- Search zones are a double-edged sword. Narrowing searches through search zones can improve results, but interacting with them adds a layer of complexity (1822).
- Users will ignore search zones until the second search (1823).

Navigation and destination webpages:

When a user searches an information environment, it’s fair to assume that he is looking for destination pages. If navigation pages are included in the retrieval process, they will just clutter up the retrieval results (1827). It’s not always clear what is a destination and navigation page – need to make careful considerations (1831). The weakness of the navigation/destination approach is that this is essentially an exact organization scheme that requires everything to be in its place (1832).

Examples:

- Homepages, the business section at the New York Times, a list of search results, the Gmail inbox, and a gallery of thumbnails – all of these pages dedicate their lives to making you go away (319).
- For example news stories, blog posts, todays weather, YouTube videos, the latest Nick Cave single, recipes, installation instructions, tutorials, wedding photos – things people have spent some energy locating and desperately want (321).

Components:

- Navigation pages exist to send the user somewhere else (design navigation pages so that they are easy to leave) (319, 320, 1826).
- Destination (or consumption) pages are the "somewhere else" you go, this is where article are read, videos watched, photos viewed and mp3s played (321, 1825).

Indexing by topic:

Indexing topics for searching makes articles and content easier to find (142).

Examples:

- In a recipe if an ingredient is mentioned, this information can be indexed to support searching by ingredient (1064).
- The Mayo Clinic employs topical search zones on its website. For example, if you're looking for a doctor to help with your rehabilitation, you might select the "Doctors & Medical Staff" search zones (1842).

Indexing for a specific audience:

Search zones can be created using the organization scheme of your website, this includes audience specific schemes (1836). If the retrieved results don't return many results for the specific audience, this method is not worth it (1841).

Example:

- The library of Michigan organizes their content by audience – librarians, libraries, and citizens (1837).

Indexing by recent content:

Chronologically organized content allows for perhaps the easiest implementation of search zones (it is a common example of search zones) (1844). Creating search zones by date is straightforward because they are exact organization schemes (1845). This might not be necessary because users who are looking within a particular date range can create an ad hoc search zone (1848).

Examples:

- The search filter of the New York Times filters by date (1846).
- Users can check for content based on filters like "past week," "Today's news," etc. (1847).

Indexing full-text:

Your first priority should be setting up your search for full-text indexing (1769).

Example:

- When searching full-text indexing, you will get results where "William Faulkner" may be mentioned, whether or not he is the author (1886).

Vertical searching (from the search miscellaneous):

Vertical search is used by a search engine that explicitly only searches within a particular subset of the vast world of documents (289). Vertical search uses topical filters to further narrow down results (291). Vertical search provides groupings of like items, which then can be sorted on like attributes (293).

Examples:

- Vertical search provides one more word to the query by adding context. For example, Chicago has a different meaning in Web search than in music search or in an airline booking search engine (290).
- Amazon.com lets users search in specific categories (vertical search) (978).

Vocabulary systems:

Vocabulary control comes in many shapes and sizes. At its most vague, it consists of any defined subset of natural language. At its simplest it is a list (2101).

Metadata:

Metadata is information about information (or data about data) and is a practical tool for information architecture (100, 736, 2096). Metadata are terms used to describe and represent content objects such as documents, people, processes, and organizations (generally an invisible IA component that affects your site) (972, 995, 2371). This is used when you organize your site; you're assigning metadata for both organization and searching/browsing (101, 559). Adding metadata to objects impact searching, browsing, filtering, and links (1241). When there are few words inherent in the content, metadata can help find it (like photos and music) (103).

Examples:

- In a song's case, it might be the following: "Brown Sugar, version 2, outtake, written by Mick Jagger and Keith Richards, performed by the rolling stones, album: Itchy fingers, bootleg, length 3:50, genre: rock and roll, blues and so on… (106).

- iStockphoto, a website with hundreds of pieces of stock photography, makes extensive use of handcrafted metadata (133).
- Bazillions does the same thing in Review Snapshot by creating categories for the tags, including pros, cons, and best uses (228).
- Index terms may be hidden as embedded metadata in an HTML document's <meta> or <title> tags. (1442, 2098).

Components:

- Metadata hidden away in source code is primarily for search engines. Dean and Deluca is telling the search engine to "crawl" their page that they sell food (117).
- Metadata can be used to generate lists of particular content (e.g., all content in the "Announcement" category) (743).
- Metadata can be used to decide what to show on a webpage (744).
- Metadata lets you generate related links based on the metadata on the content page (897).
- Database oriented (or bottom up) structures is defined by the metadata and deep contextual links embedded in the content (the photos) it contains. It allows us to apply the structure and power of relational databases to the unstructured environments of websites (1015, 1239).
- Embedded metadata is information that is extracted from the content (e.g., in a recipe, if an ingredient is mentioned, this information can be indexed to support searching by ingredient) (1063).
- Index terms are also used to make browsing easier: the metadata from a collection of documents can serve as the source of browsable lists or menus (1433).
- Metadata fields (like title, author, etc.) can be used to cluster search results for users (1988).
- When adding metadata you can choose to add metadata globally (to everything in your site), locally (to only part of the content subsection), and others only associated with a particular document type (2389).

Structural metadata:

Structural metadata is what describes the compositions or structure of an object (is it a JPEG, a 20kb file or a zip file?) (107, 737, 2372).

Examples:

- It's a JPEG, one of the most popular picture formats on the web (120).
- It's 303.29 kilobytes, which isn't terrible large (121).
- It's 609x760 pixels, which is about the size of a piece of paper (122).

Descriptive metadata:

Descriptive metadata describes what the object is about (the nature of it). This is the most commonly used metadata type used on the web (109). When you craft descriptive metadata you draw upon the stories people tell about an object (this is what people remember) (131). There should be at least a dozen ways of describing an object; this is what is used to create descriptive metadata (2409).

Examples:

- A guy selling hotdogs in New York City might contain descriptive metadata items like "food vendors," "lower west side," etc. (124, 132).
- Next to each photo, iStockphoto displays a long list of keyword links to all the photos that have been marked with those same keywords (134).
- Descriptive metadata is created by users tagging information in Etsy (227).
- Title (descriptive), Category (descriptive), Tags (descriptive) (742).

Administrative metadata:

Administrative metadata is about how an object relates to the business context (108, 2374). Administrative metadata includes not only the author/creator of the information, but the date created, the date published, and so on – everything about how the item/information was managed (123).

Examples:

- You might remember that it was taken by someone named Beatrice Abbot. Or that is was taken in 1936 (123).
- Author (administrative), Date posted (administrative), URL (administrative), Status: published (administrative) (741).

Controlled vocabulary:

A controlled vocabulary is a way to control the meaning of the vocabulary used in a website, as well as a way to keep track of related terms (145, 2453). These systems allow you to structure and map languages so that people can more easily find information (2380). There are many different kinds of controlled vocabularies (147). They change as often as websites (so they are constantly evolving) (207). You can also use controlled vocabularies to define terms that misalign with your website intent (2461).

Example:

- On a recipe site, it's good to know that salmorejo and gazpacho is essentially the same thing, just as on business site it's good to know that IBM and International Business Machines are the same in order to make sure that when people do a search, they always find something (253).

Components:

- Controlled vocabularies can be applied across a website both on the surface and for particular sections (1255).
- These include variant spellings (e.g., American or British) (2455).
- Tone (e.g., Submit or Send) (2456).
- Scientific and popular terms (e.g., cockroaches or Periplaneta Americana) (2457).
- Insider and outsider terms (e.g., what we say at work; what we say in public) (2458).
- Acceptable synonyms (e.g., automobile, car, auto, or vehicle) (2459).
- Acceptable acronyms (e.g., General Electric, GE, or G.E.) (2460).

Pros and cons:

- Creating a controlled vocabulary is subjective (166).
- Building a controlled vocabulary is hard work (185).
- Preexisting controlled vocabularies are publically available can be broadly used (1500).

Synonym rings:

A synonym ring connects a set of words that are defined as equivalent for the purposes of search retrieval (2110). The words mapped in a synonym ring are not always true synonyms (2111). When a user enters a word in the search engine, that word is checked against your synonym ring (this can be a simple text file with the list of terms) and if found it includes all synonyms (2113, 2114).

Examples:

- Flickr doesn't have a synonym ring, if you search for "photos," you wont get items tagged for "photographs" or "photo" (239).
- A user buying a food processor may enter "blender" or one of several product names (or common misspellings) (2112).
- At Frys.com a search for "iTouch" will not return results for "iPod touch" (when you don't use synonym rings the search function is limited for jargon terms) (2115).

Pros and cons:

- Problems with synonym rings is that this is happening behind the scenes, so users may be surprised about results that do not include their original search term (2117).
- Synonym rings can return less relevant results (less precision) (2118, 2125).
- Synonym rings improve recall (one study says from 20% to 80%) (2123).
- You can use synonym rings by default but order keyword matches at the top (2127).
- Or you might use synonym rings but provide options for search revision (especially if because of the synonym ring, no results are returned) (2129, 2130).
- Synonym rings are simple and useful, if you can use them, do (2131).

Authority files:

An authority file is a list of preferred terms or acceptable values. It does not include variants or synonyms. These are often used by libraries and government agencies (2132). An authority file can be a useful tool for content authors and indexers, enabling them to use the approved terms efficiently and consistently (2141).

Examples:

- The two-letter codes that constitute the standard abbreviations for U.S. states as defined by the US Postal service provide an instructive example (2136).
- Drugstor.com provides an authority file, when users search "Tilenol" in the index they are directed to "Tylenol" (2144).
- Users looking for "Tylenol" on the US Federal Drug Administration website are guided to the generic term "acetaminophen" (2149).

Classification scheme:

We use classification schemes (also called taxonomies) as an arrangement of preferred terms. It's important to recognize that these arrangements can take different shapes and serve multiple purposes (2152). Classification schemes can be used on both the front and back end of a site (2159).

Examples:

- Organize preferred terms in to groups. Rock, Hip-hop, Rap, and Techno are alike. Jazz, Bebop, and Fusion belong together (somehow) (193).
- Etsy website uses tagging and categories in their classification scheme to find fun gifts (224).
- The Dewy Decimal Classification (DDC) is a hierarchy listing that begins with top categories and gets more specific as it moves down (2155, 2466).
- Netflix has its own classification scheme for defining the different types of movie genres ("Drama," "Comedy," "Based on Real Life," "With a strong female lead," etc.) (2156).

To create their classification scheme, Netflix uses micro tags to define movie features (they are attached to each movie) (2157).

- Classification schemes can be used on searching. Wal-Mart defines its "Departments", showing their classification scheme (2158).
- The Library of Congress classification scheme was developed so that each book in a library could be placed (and found) in one and only one location (2321).
- Common examples of taxonomies include: The scientific classification for plants, animals, minerals, and other organisms (2465).

Components:

- A fronted browsable hierarch that's visible, integral part of the website (2153).
- A backend tool used for organizing and tagging documents (2154).
- There are publically available classification schemes you can use (See DDC example above) (2368).

Thesauri:

If you arrange each word from your favorite book by gathering similarly defined words, you have a thesaurus, not your favorite book (2452). It's often not obvious when a site is using a thesaurus, when it's well integrated it's invisible to the user (2199). Thesauri in websites share their history with the book version (2160).

Examples:

- The Educational Resource Information Center (ERIC) Thesaurus was designed to describe education; you can use thesauruses like these to help with labeling problems (1502, 1504).
- PubMed will be used an example of good thesaurus. It has a huge thesaurus with over 19,000 terms to leverage search (2200).
- If a user is unsure whether to use the term "tropical storm" or "hurricane," accessing a thesaurus can identify the preferred term (2412).

Component:

- A thesaurus on the backend can enable a more seamless and satisfying user experience on the frontend (2095).

Classic thesaurus:

Classic thesauri are used on the web not to just get better words, but also to create an interconnected web of words to help people find things they don't have (161). The classic thesaurus connects synonyms, homonyms, antonyms, broader narrower terms, and related terms (2161). They do this so that the ambiguity of language doesn't stop people from finding things (2163). One of the big advantages of using a classic thesaurus is the power and flexibility to shape and refine the user interface over time (2209).

Examples:

- A classic thesaurus says, "gravlax is a type of salmon that is the same as cured salmon and is an ingredient for bagels and lox" (149).
- MeSH browser uses a thesaurus, so when searching for one term, the preferred term will always be returned (2205).

Indexing thesaurus:

Indexers use the thesaurus to map variant terms to preferred terms when performing document-level indexing (2218).

Components:

- Indexing thesauri promote consistency and flexibility. Indexers work together to gain a shared understanding of preferred terms and indexing guidelines (2230).
- Indexing thesaurus allows you to build browsable indexes of preferred term, enabling users to find all the documents about a particular subject or product through a single point of access (2231).

- Indexing thesauri are useful when used over time (like intranets), where users can begin to recognize the preferred terms (2232).

Searching thesaurus:

Searchers use thesaurus for retrieval, whether or not they're aware of the role it plays in their search experience (2219). They can use a search thesaurus to narrow down search results, it can provide greater browsing flexibility, and search thesauri can become a portal to provide new ways to navigate and access large amounts of information (2254, 2256, 2262). Searching thesauri are cheaper to develop and maintain (in relation to the amount of content on your site) (2263).

Example:

- For example, when a user enters a term into the search engine, a searching thesaurus can map that term (2242).

Thesaurus standards:

There are many different existing thesaurus standards that you can use in your website (rather then reinventing the wheel) (187).

Examples:

- The ANSI/NISO standard is entitled "Guidelines for the construction, Format, and Management of Monolingual Thesauri" (2267).
- The ANSI/NISO standard presents simple guidelines that are difficult to follow. Standards do not mean that that you don't have to put in some work (2268).

Pros and cons:

- Standards involve good thinking baked into guidelines (2271).
- Most thesauri management software works with established standards (2272).
- Compliance with a standard will provide better database compatibility (2273).
- Read the guidelines, follow the standards, but prepare to deviate when necessary (2274).

Thesaurus terms:

It's useful to know the thesaurus terms that define the connections made between terms (2181). The preferred terms are a tool to control vocabulary and keep everyone on the same page, as well as to inform your labeling process (189, 2143). The preferred term is the center of the thesaurus terminology universe, but one person's preferred term can be another's related term (2193). As your content grows you need to define your terms, you can't have hundreds of hits for every preferred term (2312).

Examples:

- Predetermined vocabularies of preferred terms describe a specific domain (e.g., auto racing or orthopedic surgery); typically include variant terms (e.g., "brewski" is a variant term for "beer") (1070).
- For example, if we're working on a website for Knowledge Management magazine, the single term "knowledge management software" or perhaps "software (knowledge management)" may be the way to go. However, if we're working on a broad IT site like CNET, it may be better to use "knowledge management" and "software" as independent preferred terms (2314).

Components:

- Preferred term or accepted term is the most important, all relationships are defined according to it (2182).
- Variant terms are loosely synonymous with preferred terms (2183).
- Broader terms are the parents of preferred terms (2184).
- Narrower terms are the children of the preferred term (2186).
- Related terms are related to the preferred term (2188).
- Use determines the preferred term. Indexers may want to use one term but according to the rules they should use another (2190).
- Used For is the reciprocal term for preferred term (shows list of variants on the preferred terms record) (2191).

- Scope notes can define the preferred term to restrict its meaning (e.g., pitch as in wood sap not pitching a ball) (2192). They can increase the specificity, deliberately restricting a meaning to one term (2309).
- Parenthetical term qualifiers provide a way to manage homographs (How you clarify if the term is pitch (tree sap) or pitch (throw ball)) (2308).

Semantic relationships:

These are the rich semantic relationships created between the vocabulary in a website (2211).

Hierarchy relationships:

The hierarchical relationship enables the classification of preferred term into categories and subcategories (relating broader and narrower concepts through the familiar parent-child relationship (1384, 2174, 2284). You need to define the number of levels of hierarchy relationships you will have on your website (granularity) (2293). In a strict hierarchy, each term appears in only one place, this was originally the plan for biology taxonomies (2315).

Examples:

- Yahoo!'s hierarchical relationships are identified by exploring concepts related to the term you searched (159).
- Perhaps you decide Hip-Hop, Rap, and Techno are all subsets of Rock. Maybe Hip-Hop is a subset of Rap. Looking at Hip-Hop and Techno, maybe you decide these are aspects of Club Music. You can start to form a hierarchy (195).

Types:

- Generic – This is the traditional class–species relationship we draw from biological taxonomies. Species B is a member of Class A and inherits the characteristics of its parent. For example, Bird NT Magpie (2286).
- Whole part – In this hierarchical relationship, B is a part of A. For example, Foot NT Big Toe (2287).

- Instance – In this case, B is an instance or example of A. This relationship often includes proper names. For example, Seas NT Mediterranean Sea (2288).

Polyhierarchical relationships:

This allows you to make information available in multiple ways (2363). If you have the potential for multiple hierarchy relationships, you have a polyhierarchical relationship (196). These are unavoidable in large information systems, most systems allows for the notion of primary and secondary locations (2318, 2323). In digital information systems, the only real challenge introduced by polyhierarchy is representing navigational context (2322).

Examples:

- iTunes uses polyhierarchical relationships between music terms/classification (200).
- For example, MEDLINE cross-lists viral pneumonia under both virus diseases and respiratory tract diseases (2319).
- Wikipedia is another large information environment that makes extensive use of polyhierarchy. At the footer of most articles in the Wikipedia website is a box with links to the higher levels in the hierarchy that list that particular article (2320).

Equivalence relationships:

The equivalence relationship is employed to connect preferred terms and their variants. While we may loosely refer to this as "synonym management" it's important to recognize that equivalence is a broader term than synonym (2278). Grouping terms together using equivalence relationships are for the purpose of search retrieval (2279). You can also fold general and specific terms in to equivalence relationships (2281).

Examples:

- Equivalence relationships: cured salmon and gravlax are the same for the purpose of a search (151).

- The relationship can be as simple as two words for the same thing: cat and kittycat. These are synonyms (152).
- They can also be different spellings or acronyms for the same thing. Lion is lyon; SPCA is Society for Prevention and Cruelty to Animals (variants) (153).
- The words can be slightly different, but for the purpose of search, you may choose to treat them the same: cat and kitten (154).

Associative relationships:

The associative relationship is often the trickiest, and by necessity is usually developed after you've made a good start on the other two relationships (2294). There is the notion that associative relationships should be "strongly implied." For example, Hammer RT Nail. In practice, however, defining these relationships is a highly subjective process (2298).

Examples:

- An associated item, Pumpernickel, which isn't cheese at all, but makes good eating (184).
- Cheese leads to crackers (202).
- A Beck CD leads to concert tickets (203).
- Model associative relationships between concepts (e.g., See Also, See Related), (2106).
- The relationship is often articulated through use of See Also. For example, Tylenol See Also Headache (2189).
- Associative relationships allow what marketing folks call "cross-selling," allowing an ecommerce site, for example, to say "Hey, nice trousers! They'd go great with this shirt." (2302).

Pros:

- The associative relationship provides an excellent vehicle for connecting customers to related products and services (2301).
- These associative relationships can both enhance the user experience and further the goals of the business (2303).

Faceted Classification:

Faceted classification uses many different facets (delimiting terms) to narrow down/browse through information to get to the content you need (78, 79, 177). When sites have multiple parents, it's called faceted classification. The facets can include any quality shared by a number of items, including price, weight, and colour; or in this case, brand origin, and firmness (176). The more facets something has the more ways it can be organized (2481).

Examples:

- Allrecipes mixes several kinds of classifications. This means some bread recipes can be in more than one category. For example, you might find a hot cross bun recipe under Breakfast Pastries, Holiday Breads, Yeast Breads, Fruit Breads, and Rolls and Buns. This is okay. It may make purists itch, but it gets people to the bread recipe they need (73).
- Let's look at equipment for a moment; what are the facets of a pan? Some common ones might be shape, material, brand, and use (82).
- Shopping often lets you filter by category, price, and other pertinent facets, depending on the item being researched, for example: megapixels for cameras or size for clothing (292).
- Epicurious allows customers to browse recipes by choosing facets that include cuisine, meal, and type of dish. These facet labels act as dividers on the browse page, provide context to the list items, and make it easier to scan (351).
- Wine.com provides a simple example of faceted classification. Wine has several facets that we commonly mix and match in our selection process at restaurants and grocery stores (2329).
- Using the record store as an example, the following facets are available for each record: Record name, Artist name, Record label, length, etc. (2482).

Pros and cons:

- If there is too little content, faceted classification will lead the users back to the same thing (88).

- The best faceted browse systems let you determine which criteria you start with and which sequence you use. You never get zero results as facets aren't displayed where they're not valid (597).
- Filters allow people to narrow down a large set of content. They are great content sets, where the content has a range of attributes and people may want to approach the site with different starting point (938).
- The nice thing about a faceted classification approach is that it provides great power and flexibility (2341).
- If a particular facet is interesting but the data to support it doesn't exist or is hard to gather, it might not be the best plan to use that facet (2483).

Sources for Heuristic Descriptions

Source 1

- Wodtke, C. & Govella, A. (2009). *Information Architecture: Blueprints for the Web* (2nd ed.). A. Govella (ed.) Thousand Oaks, CA: New Riders Publishing.
- Codes 1-579

Source 2

- Spencer, D. (2011). *A Practical Guide to Information Architecture*. Five Simple Steps.
- Codes 580-963

Source 3

- Rosenfeld, L., Morville, P. & Arango, J. (2015). *Information architecture: For the web and beyond* (4th ed.). Sebastopol, CA: O'Reilly Media Inc.
- Codes 964-2444

Source 4

- Covert, A. (2014). *How to Make Sense of Any Mess: Information Architecture for Everybody*. CreateSpace.
- Codes 2445-2516

Appendix L – Digital Museum Heuristic Evaluation Guide

Contextual framework:

What does the website say about:

- The business mission/model (excerpt and summary) – Context
- The user/audience (who, explanation, examples) – Audience
- The website content (what, where, explanation, examples) – Content

Organization systems:

Organization schemes:

- What types of schemes does the website have?
- Where are they located?
- How do the websites utilize the schemes they've chosen?

Organization structure:

- What structure type do they have?
- How does the site utilize their structure?
    - E.g., where in the site, how, when, etc.

Include examples and clarifying information

Sketch/diagram the digital museum structure

Labeling system:

Types of labels:

- What types of labels do the sites have?
    - Contextual label examples
    - Headings examples

- Navigation label examples
- Index term examples
- Icon label examples

- Additional thoughts/notes on the labels

Labeling consistency:

- Explain consistency issues found in the digital museum labels
- Consider components of consistency
    - Examples (good and bad)

Include examples and clarifying information

Navigation Systems:

What types of navigation does the site have?

- Where, why, when, how, what features, etc.
- Examples

Are there any issues with the navigation?

Include examples and clarifying information

Search Systems:

Look at the search system components:

- Can you tell anything about the search algorithm?
- What is the recall and precision?
- What does the search interface look like?
    - Placement, components, etc.
- Does it have advanced search? Describe it
- How are search results displayed?
    - Format, information displayed, etc.

- How are the search results sorted?
- How are the search results ranked?
- Does the search system use best bets?
- Are there any additional actions available after searching?
    - Save a search
    - Select a subset of results
    - Narrowing results down
    - Repeating/new search
- What query builders are available?
    - Spell checkers
    - Phonetic tools
    - Stemming
    - Natural language processing tools
    - Autocomplete/autosuggest
- What query language(s) is supported?
- What content has been indexed for searching?
- Does the site have search zones?
    - What are they?
    - How are they used?
    - Does it have vertical search?

Include examples and clarifying information

Vocabulary systems:

Metadata:

- What metadata is present?
    - Structural metadata?
    - Descriptive metadata?
    - Administrative metadata?
- Where and how is the metadata used in the site?

Include examples and clarifying information

Controlled vocabulary:

- Does the site have a controlled vocabulary?
- What components of a controlled vocabulary are present?
- Does it use a synonym ring? (How, what, where, etc.)
- Does it use an authority file? (How, what, where, etc.)
- What classification scheme(s) is present on the site? (How, what, where, etc.)

Include examples and clarifying information

Thesaurus:

What type of thesaurus does the digital museum have? Does it have a thesaurus?

- Components?
- What are the preferred terms?
- How is the thesaurus structured?
- Does it follow standards? (Existing or created)

Look at indexes (if available) and descriptive metadata to determine the thesaurus structure/preferred terms.

Include examples and clarifying information

Semantic relationships:

- How does the site use hierarchical relationships?
- How does the site use equivalence relationships?
- How does the site use associative relationships?

Include examples and clarifying information

Faceted classification:

Does it use faceted classification?

- What are the facets?
- Where is this feature available?

Include examples and clarifying information

Appendix M – Example Museum Report

National Portrait Gallery

Contextual Framework

Context

In order to continue to successfully interest their audience, they must embark "…on a major renewal programme designed to transform the services it provides to its visitors, physical and virtual, and in the way it presents and interprets its collection. The renewed National Portrait Gallery will deepen understanding and enjoyment of its remarkable collection of portraits and will broaden its appeal to the widest possible audience." (National Portrait Gallery Corporate Plan 2016-2019, 2016, Our ten year vision).

The National Portrait Gallery also has a digital strategy, that outlines how improvements will be made to their website going forward.

"The National Portrait Gallery has three clear transformational aims in terms of its future strategy to be realized through the *Inspiring People* project and these translate directly into its digital aims:

- Access – to enable use of its collections and related content.
- Understanding – to encourage participation and engagement.
- Sustainability – to ensure revenue streams to support its activities"

(Digital Strategy, Nov. 2016, Digital Aims)

Audience

The website is only available in English, however, it does provide visitor guides in ten different languages (French, German, Italian, Spanish, Polish, Russian, Japanese, Simplified Chinese, Traditional Chinese, and Arabic). This tells me that though the website is English (its main audience or it assumes that most users can speak some English) it is trying to reach an international audience. These visitor guides contain information about the physical National Portrait Gallery location (hours, gift shop, accessibility, toilets, entrances, etc.) as well as an introduction to the gallery and its content. The NPG shop also provides currency conversion from GBP to either the Euro and USD (more international impact).

The National Portrait Gallery organizes some of its content by audience – Adults, Schools and Colleges, Families, Young People, Corporate Support, etc. By organizing content around these audiences, the website claims them as their "typical" user. Content available for these audiences include workshops, resources (broken down by grade level), special events, projects for 14-21s, partnership programs, etc.

The National Portrait Gallery also provides a rich area of resources for researchers. This includes researching programs, access to the archives and library, past research work on portraits, family history information and much more. The inclusion of these materials mark that the "researcher" is an important NPG audience member.

The National Portrait Gallery also makes accessible content available throughout the site, for both the physical gallery (e.g., wheelchair accessible points of access) and online (e.g., large print guides and options). The NPG does a good job at trying to make all their content accessible to every user.

Content

The NPG contains information about:

- The physical gallery – hours, directions, floor plans, shop, restaurant, audio guides, cloakroom, etc.
- Events happening at the NPG – current exhibits, future exhibits, past exhibits, etc.
- Collections of portraiture (paintings, drawings, sculpture, etc.).

- Learning resources for specific audiences – workshops, lectures, holiday events for families, etc.
- Information about becoming a member and joining the National Portrait Gallery (for both individuals and corporations).
- Items for sale (under the subsite National Portrait Gallery shop).

Organization System

Organization schemes

The National Portrait Gallery uses a hybrid organization scheme (a mixture of both exact and ambiguous organization schemes).

The top-level schemes:

- Topical: Home, What's On, and Collections
- Task: Visit, Learning, Join & Support, and Shop

Second level schemes (excerpted):

- Format schemes: Audio guides, Books, Portrait Prints, etc.
- Chronological: Current Exhibitions, Today's Display, Today's Events, Future, and Past.
- Topical: Opening hours, Portrait explorer, Event programming, Primary Collections, Reference Collections, Resources, £10m Portrait Fund Challenge, Trusts and Foundations, Themes, Just Arrived, etc.
- Task: Eat. Drink. Shop., Visit the Collection, Plan your visit, Search the Collections, Explore further, Join Us, Contact Us, etc.
- Audience specific: Adults, Group Visits, Family Connections, Schools and Colleges, Families, Young People, For Members, Customer Service, etc.

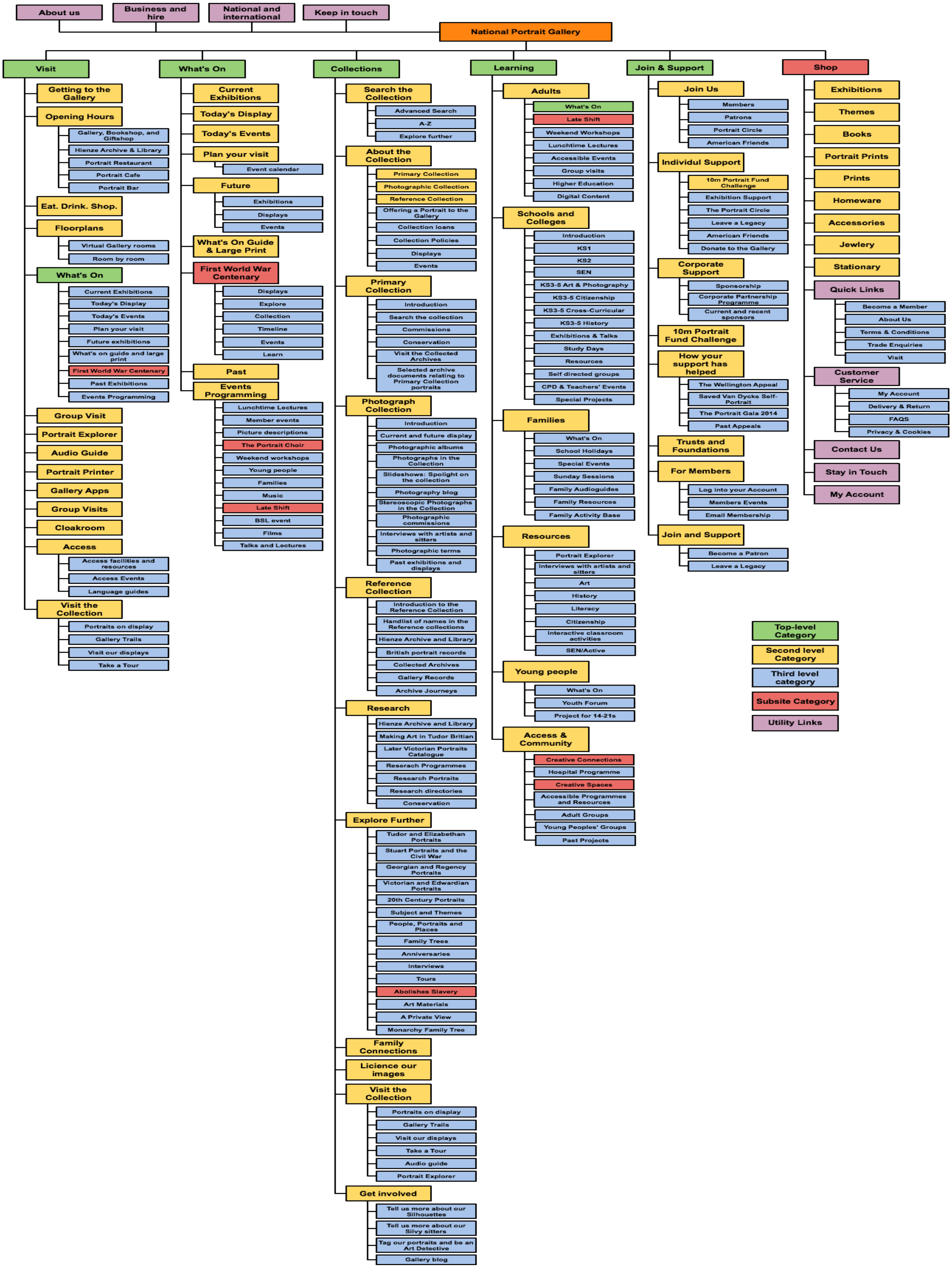
Sub-level organization:

There are many different types of organization schemes within the sublevels of the NPG.

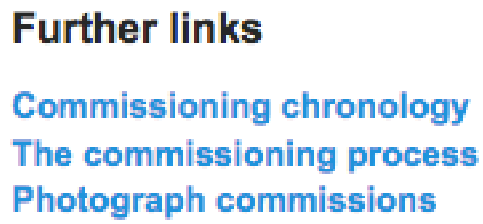
- **Alphabetical**: there are multiple A-Z indexes (for example, the hand list of names in the Reference collection, Photographic terms, Artist suppliers, Restorers, etc.), and it is used as a secondary organization scheme (e.g., tours, organized by topic are organized alphabetically).
- **Chronologically**: This is the main exact organization scheme used in the lower levels. For example, past exhibits and displays (in multiple year section, e.g., 2011-2015), Take a Tour/Tour organizes information by time period (e.g., Regency and 20th Century), the NPG blog organizes the posts from newest to oldest, Events and Exhibitions are organized chronologically (like in the Event Calendar), etc.
- **Geographically**: This is not used that often, but in the subpage "People, Portraits, and Places" is organized by regions (UK, Europe, the Americans, etc.).
- **Format**: Under Research directories you can view a list (alphabetically organized) of artists materials and what portraits utilize them. Information is also organized under Slideshows, Audio guides, Interviews, etc. Additionally, this organization scheme is found within subsites (e.g., Late Night has a Film category).
- **Organizational (business) scheme** is used under the "Business and hire" category (where the corporate aims, procedures, etc. are found).
- **Topical** is used throughout the site. For example, Tours is organize by themes, "Subject and Themes" page (which organizes the five categories that represent the keywords the content is tagged with), along with most subheadings in categories.
- **Task-oriented**: This becomes less common the further into the polyhierarchy you move, however you still have options like "Visit," "Learn" (Common in subsections like First World War Centenary), "Meet the Team," etc.
- **Audience specific** organization schemes are most common under the "Learning" collection (e.g., Youth Forum, KS2 – ore Kindergarten kids aged 3-6) as well as in the Join & Support category (e.g. "Members," "Corporate Partnership programme," etc.).

Organization Structure

The NPG uses a Hybrid Organization structure, with a poly hierarchical top-level structure and subsites scattered within the structure. A database-oriented structure is used to organize the collection catalogue and the archives and library catalogue.


About us
Business and hire
National and international
Keep in touch
National Portrait Gallery
Visit
Getting to the Gallery
Opening Hours
Gallery, Bookshop, and Giftshop
Hienze Archive & Library
Portrait Restaurant
Portrait Cafe
Portrait Bar
Eat. Drink. Shop.
Floorplans
Virtual Gallery rooms
Room by room
What's On
Current Exhibitions
Today's Display
Today's Events
Plan your visit
Future exhibitions
What's on guide and large print
First World War Centenary
Past Exhibitions
Events Programming
Group Visit
Portrait Explorer
Audio Guide
Portrait Printer
Gallery Apps
Group Visits
Cloakroom
Access
Access facilities and resources
Access Events
Language guides
Visit the Collection
Portraits on display
Gallery Trails
Visit our displays
Take a Tour
What's On
Current Exhibitions
Today's Display
Today's Events
Plan your visit
Event calendar
Future
Exhibitions
Displays
Events
What's On Guide & Large Print
First World War Centenary
Displays
Explore
Collection
Timeline
Events
Learn
Past
Events Programming
Lunchtime Lectures
Member events
Picture descriptions
The Portrait Choir
Weekend workshops
Young people
Families
Music
Late Shift
BSL event
Films
Talks and Lectures
Collections
Search the Collection
Advanced Search
A-Z
Explore further
About the Collection
Primary Collection
Photographic Collection
Reference Collection
Offering a Portrait to the Gallery
Collection loans
Collection Policies
Displays
Events
Primary Collection
Introduction
Search the collection
Commissions
Conservation
Visit the Collected Archives
Selected archive documents relating to Primary Collection portraits
Photograph Collection
Introduction
Current and future display
Photographic albums
Photographs in the Collection
Slideshows: Spotlight on the collection
Photography blog
Stereoscopic Photographs in the Collection
Photographic commissions
Interviews with artists and sitters
Photographic terms
Past exhibitions and displays
Reference Collection
Introduction to the Reference Collection
Handlist of names in the Reference collections
Hienze Archive and Library
British portrait records
Collected Archives
Gallery Records
Archive Journeys
Research
Hienze Archive and Library
Making Art in Tudor Britian
Later Victorian Portraits Catalogue
Research Programmes
Research Portraits
Research directories
Conservation
Explore Further
Tudor and Elizabethan Portraits
Stuart Portraits and the Civil War
Georgian and Regency Portraits
Victorian and Edwardian Portraits
20th Century Portraits
Subject and Themes
People, Portraits and Places
Family Trees
Anniversaries
Interviews
Tours
Abolishes Slavery
Art Materials
A Private View
Monarchy Family Tree
Family Connections
Licience our images
Visit the Collection
Portraits on display
Gallery Trails
Visit our displays
Take a Tour
Audio guide
Portrait Explorer
Get involved
Tell us more about our Silhouettes
Tell us more about our Silvy sitters
Tag our portraits and be an Art Detective
Gallery blog
Learning
Adults
What's On
Late Shift
Weekend Workshops
Lunchtime Lectures
Accessible Events
Group visits
Higher Education
Digital Content
Schools and Colleges
Introduction
KS1
KS2
SEN
KS3-5 Art & Photography
KS3-5 Citizenship
KS3-5 Cross-Curricular
KS3-5 History
Exhibitions & Talks
Study Days
Resources
Self directed groups
CPD & Teachers' Events
Special Projects
Families
What's On
School Holidays
Special Events
Sunday Sessions
Family Audioguides
Family Resources
Family Activity Base
Resources
Portrait Explorer
Interviews with artists and sitters
Art
History
Literacy
Citizenship
Interactive classroom activities
SEN/Active
Young people
What's On
Youth Forum
Project for 14-21s
Access & Community
Creative Connections
Hospital Programme
Creative Spaces
Accessible Programmes and Resources
Adult Groups
Young Peoples' Groups
Past Projects
Join & Support
Join Us
Members
Patrons
Portrait Circle
American Friends
Individul Support
10m Portrait Fund Challenge
Exhibition Support
The Portrait Circle
Leave a Legacy
American Friends
Donate to the Gallery
Corporate Support
Sponsorship
Corporate Partnership Programme
Current and recent sponsors
10m Portrait Fund Challenge
How your support has helped
The Wellington Appeal
Saved Van Dycks Self-Portrait
The Portrait Gala 2014
Past Appeals
Trusts and Foundations
For Members
Log into your Account
Members Events
Email Membership
Join and Support
Become a Patron
Leave a Legacy
Shop
Exhibitions
Themes
Books
Portrait Prints
Prints
Homeware
Accessories
Jewlery
Stationary
Quick Links
Become a Member
About Us
Terms & Conditions
Trade Enquiries
Visit
Customer Service
My Account
Delivery & Return
FAQS
Privacy & Cookies
Contact Us
Stay in Touch
My Account
Top-level Category
Second level Category
Third level category
Subsite Category
Utility Links

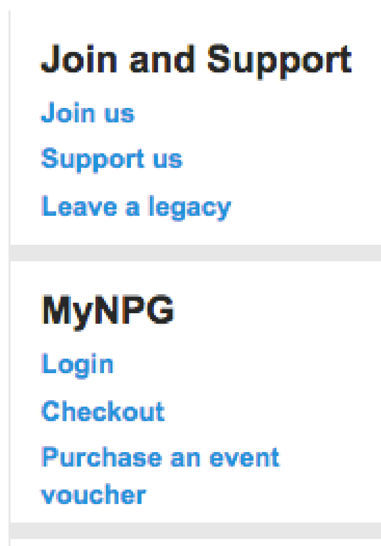
This site is polyhierarchical because much of the content is cross-listed under different categories and at different levels.



Examples:

- You can access the "What's On" top-level category from the Global navigation bar, under "Visit" top-level category and under "Learning" > "Adults."
- "Resources is both a second level category under "Learning" and used as a third level category under "Schools and Colleges."
- The subsite "First World War Centenary" is organized both under "What's On" (as a second level category), "Group Visits," (third-level), and under "Collections" > "Explore Further" > "20th Century portraits" > "The Great War in Portraits" (fourth level).

There are multiple subsites through the NPG site; these take you to a new URL (though they maintain the NPG logo) that has additional organization schemes (e.g., Picture the Poet has a hierarchy structure). These tend to be simply organized and straightforward, but are still considered part of the NPG website (you use the NPG logo to navigate back to the NPG homepage).



The database-oriented structure is only viewable in the catalogue/holdings (within the portrait collection and the archives and library). The items in the collection contain a title, creator, publication (if applicable), date, call number, etc. The items in the Archive and Library catalogue contain a title, publisher, location, call number, and related website.

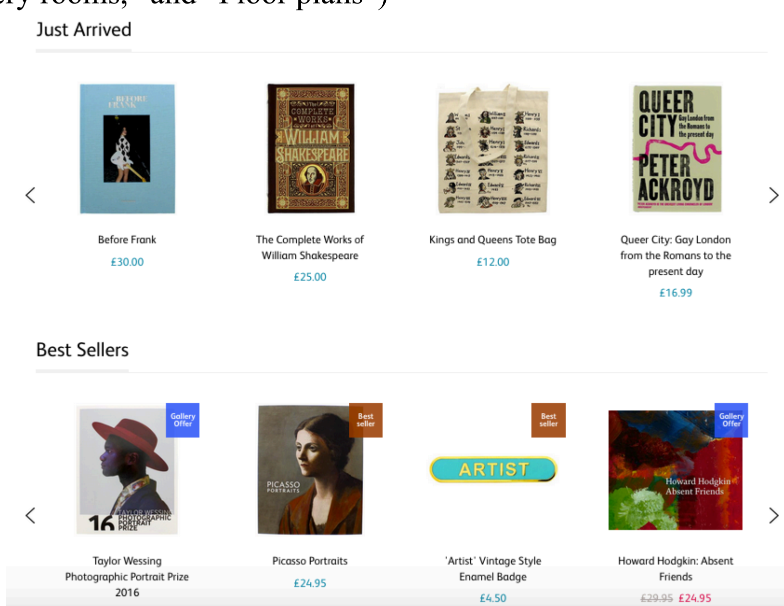
Labeling Systems

Types of Labels

The NPG has all types of labels found throughout their site.

Contextual labels: These are both found organized under "Related" or "Further Links" and within the text. Additionally there are contextual sections at the bottom of select pages (with images, title, description and "Find out more" links).

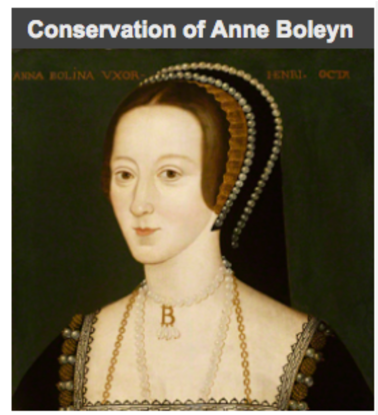
- For example, the "Room by room" page has in-text links (to different rooms in the physical NPG) and sections for related pages at the bottom of the page ("Events calendar," "Virtual Gallery rooms," and "Floor plans")



- Contextual links to "Related sitters" or "Related artists" is available in the "Handlist of names in the Reference Collection."
- "Further links" contextual links can be found in the "Commissioning portraits" page – offering similar pages of interest.
- Problem: the contextual links are not consistently applied across the site (except "Join & Support" links found on every page on the right side (relevant or not)).

Headings are throughout the site, as page titles, subsection titles amongst others.

- For example, in the big footer at the bottom of the page, headings organize the persistent links (“About us,” “Business and hire,” “National and international,” “Keep in touch”).
- Headings are also used to organize/provide context to contextual links, e.g., “Other ways to support the Gallery” organizes links to “Exhibition Supporter Groups” and “Portrait Circle” (which may not necessarily seem to relate).
- Headings are usefully used to tell users what has just arrived in the NPG shop as well as the shops “Best Sellers.”



Navigational labels are used throughout the site.

- Global, local, breadcrumb, utility, indexes, etc.
- The navigation labels cover all aspects of the site (though there are some inconsistency issues, see below).
- For example, within the “Opening hours and admission” page there are navigation labels for each section that lists its opening hours. These links take the user to the digital version of that physical location.
    - Gallery, Bookshop and Gift shop > Eat. Drink. Shop.
    - Hienze Archive & Library > Hienze Archive & Library homepage
    - Portrait Restaurant > Portrait Restaurant and Bar
    - Portrait Café > Eat. Drink. Shop.
    - Portrait Bar > Portrait Restaurant and Bar
- Many of the navigation labels (local) are combined with images – to provide context and introduce the audience to their collection (see image – this tells the user where they are going and what Anne Boleyn looked like).

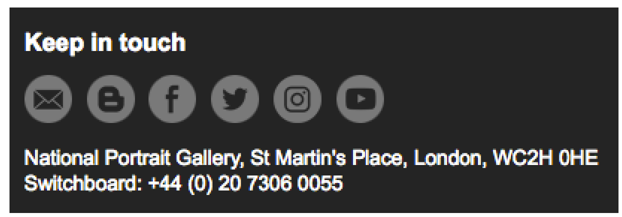
NPG uses index terms (keyword, metadata, tags, etc.) throughout the site to provide more information about the content.

- There is an entire page called “Subjects and Themes,” which indexes the five main keywords/themes that the portraits are indexed with for browsing.

- The descriptive metadata included with collection items (when searching) includes these subjects and themes as well as descriptions.

Icon labels are used in the main NPG site as well as the "Shop" (there is an icon shopping basket and a flag, which represents the currency selected).

- The icon labels in the main site represent social media pages (Facebook, Twitter, Instagram, YouTube), email, and the NPG blog, as well as search (magnify glass icon) and even more connection/social media sites.



Labeling Consistency

NPG labeling inconsistency:

- Style: capitalization is an issue, some second words are capitalized and others are not. For example, "BSL events" and "Members' events" does not capitalize events, but "Today's Events," "Events," and "Events Calendar" are capitalized. "About us" in the main page does not capitalize the "U," but "About Us" in the shop does.
- Inconsistent labeling of sub pages. For example, "Portraits on display" leads to the "Room by room" page, but is called something different. This gets even more complicated because "Portraits on display" within the "Tudor section" leads to a subsection within the "Room on room" page (inconsistent linkage, though linking to the specific Tudor pages does make sense). Additionally, all these pages "Lunchtime lectures," "Member events," "Picture descriptions," Weekend Workshops," Young people," and "Families" all lead to the event calendar, but have different titles.
- There are some terminology/jargon issues under "School and Colleges." For example, "KS1," "KS2," "KS3," etc. are all categories that describe different age groups/school levels and as a non-British person this was hard to initially decipher (without reading the text of the pages), however I assume that British people would understand this terminology.
- Labeling terminology: There are links that lead to the same page, but have slightly different labels. For example, "Take a Tours" and "Tours" lead to the same page.

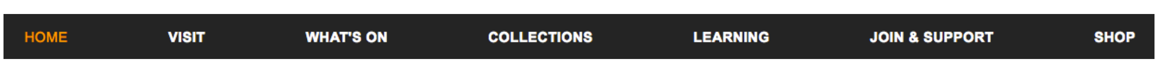
- Granularity - items are cross-listed with labels, which appear as the top-level and in sub levels, but they have the same label. For example, “About the Collection” contains categories which are second-level categories (under Collections) like “Primary Collection,” “Photograph Collection,” and “Reference Collection.” “Group Visits” is a second-level category under “Visit” and also a third level category under “Learning” > “Adults.”
- Use of the same label for different pages - What's On is a top-level category, but it is also indexed within “Learning” > “Adults.” Additionally, the label “What’s On” is organized under “Learning” > “Families” and “Learning” > “Young people” but they load different pages when selected (though they do still have information about events).

Navigation Systems

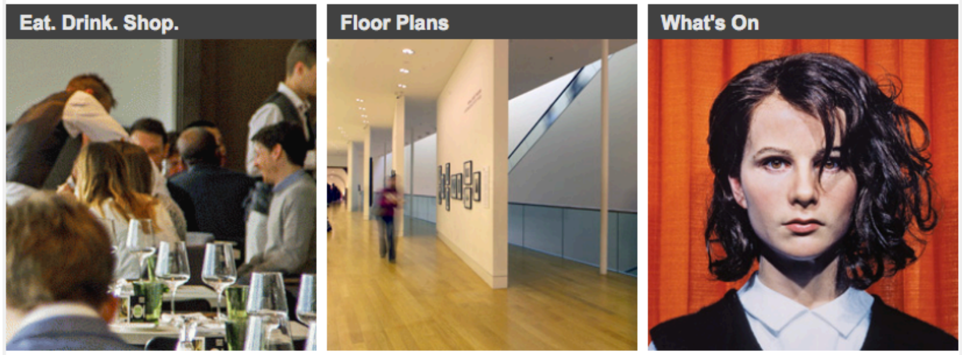
The NPG has examples of every type of main navigation type (Global, Local, contextual/hypertext, breadcrumb, and utility navigation).

The Global navigation toolbar (at the top of the page) is simple; it has a home page link along with six other categories (Visit, What's On, Collections, Learning, Join & Support, and Shop). This toolbar does not have an expandable menu to see the second level of categories.

- The subsites have different global navigation, but they are all structured similarly. These show the secondary level organized underneath the top-level category when the top-level is selected.



The Local navigation is structured in two ways. The first local navigation option brings the user to the homepage for that category; this is a page of artfully arranged links, which combined portrait images, description text (that appears when the use hovers their cursor over the option), and/or additional links (see image).



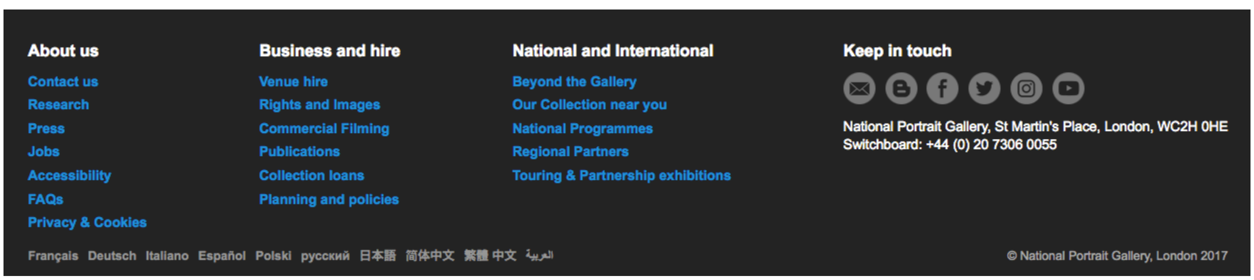
The second type of local navigation is within lower levels (particularly within the sub levels in the Learning section) and is organized into three columns, populated with the sub categories (the towards the top of the page below a secondary banner). Selecting these options only changes the information in the text field below the columns, unless it navigates to a new subsection (this doesn't happen that often).

Contextual navigation links are found throughout the site. These can be organized under "Related" or "Further Links" and within the text. Additionally there are contextual sections at the bottom of select pages (with images, title, description and "Find out more" links). See contextual labels for more information.

Breadcrumb navigation is present in the site, but there are some issues with it. The polyhierarchical structure of the website means you can access subpages and content pages in multiple sections, however the breadcrumb navigation shows the path closest to a top-level category. For example when you access "Portrait explorer" within the "Learning" category the breadcrumb navigation should be "Learning / Resources / Portrait explorer," but when you click on it the breadcrumb navigation shows "Home / Visit / Portrait Explorer." The breadcrumb navigation is located at the top of the content page (below the banner). Another example is that when selecting "What's On"

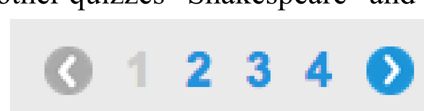
links under visit gives the user the breadcrumb navigation “Home/What’s On” not “Home/Visit/What’s On.”

Utility navigation is found at the bottom of the page in a large footer as well as a few options that persist at the top of the page. The utility navigation options include links to the NPG blog, social media, information about the NPG, Business and hire, and National and international, and links to other languages (visitor guides). The top of the page has the NPG logo (a link to the homepage), a search box and a link to even more social media sites.



Supplementation navigation features:

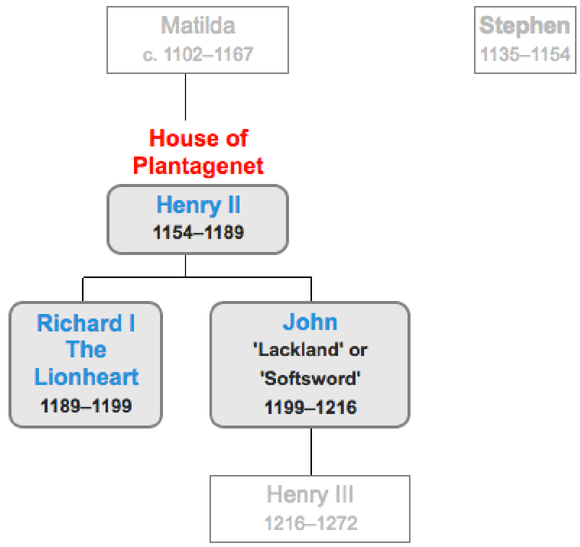
- The NPG has multiple A-Z indexes. For example, the handlist of names in the Reference collection, Photographic terms, Artist suppliers, Restorers, etc.
- There are numerous guides throughout the site as well. For example, the “Who do you think you were?” is a quiz that tells you who you would be in Elizabethan era. This guides you through the quiz. Additional guides include other quizzes “Shakespeare” and “Votes for Women,” as well as “History of Hair and Beauty,” which guides users through the beauty steps of different eras.



- There are “Control Panels” for when you join as a member (the form is all one page).
- Pagination navigation is used on the main pages (to scroll through highlighted elements) as well as in the search fields (for the collection and archives and library catalogues).

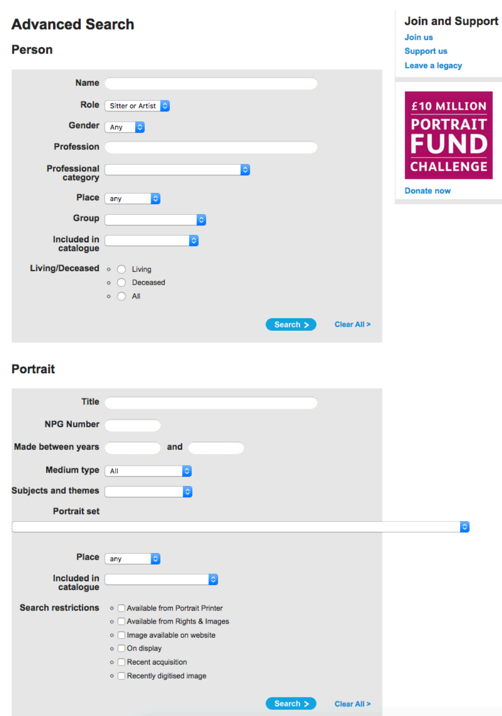
- This website does not have a sitemap or toolbars.

Advanced navigation features:



- The NPG uses visualization features for navigation. This includes maps and timelines. For example, under the Tudor and Elizabethan subpage there is a navigational timeline. There is also family tree navigation, where you can click on the names within the structure to see their portraits.
- Social navigation, the NPG gives users the option to become "Art detectives" (if they sign up), they can then tag the portraits with keywords. Additionally the NPG blog allows users to navigate the content with pre-existing tags.

Search Systems

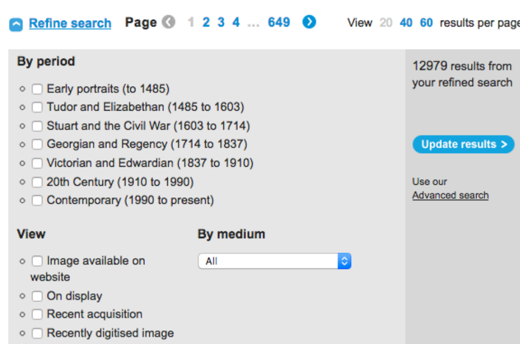


The National Portrait gallery has numerous search systems throughout their site and subsite.

Main/Collection catalogue search:

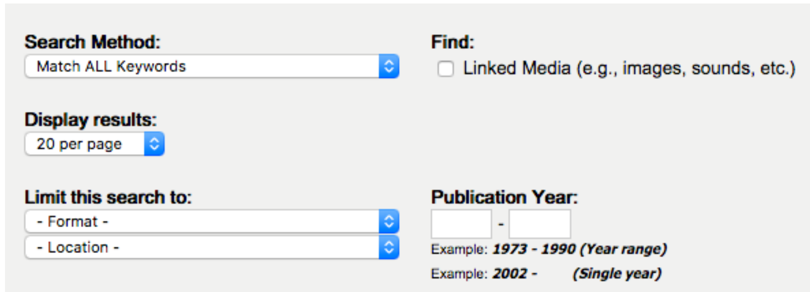
- This search system is both the system found in the utility navigation as well as the system that searches the collection catalogue.
- The algorithm is unknown
- The search seems to prioritize precision over recall. A search for "Earl of Sandwich" brought up 68 results and they all either had an "Earl of Sandwich" (I, II, etc.) in the title or in the "Sitter field."

- The search interface is found in multiple locations (at the top of the page, in a navigation section within the top-level Collections category, at the top of the search results). It is composed of a search box (with rounded corners) and a magnify glass icon for the button (or a right pointing arrow for the version at the top of the search results and in the collections page).
- It does have an advanced search feature. There are two sections that the user can add more narrowing information to – "Person" and "Portrait." The features include selecting a role (artist or sitter), entering a profession, selecting a professional category, living/deceased, the NPG number (associated with each portrait), subject and themes associated with it, etc.
- The search results are displayed in a list format with 20 items per page (though that can be expanded to view 40, 60 results per page), the user can choose to view either as a list or in thumbnails. The information displayed for each search result differs depending on what items you were viewing (People, Portraits, Events, and Other pages). As well as if you are viewing them as thumbnails (then you have to role over the images to see information).
    - People display – the title/person name, the date, the profession/role, and how many portraits the sitter is associated with.

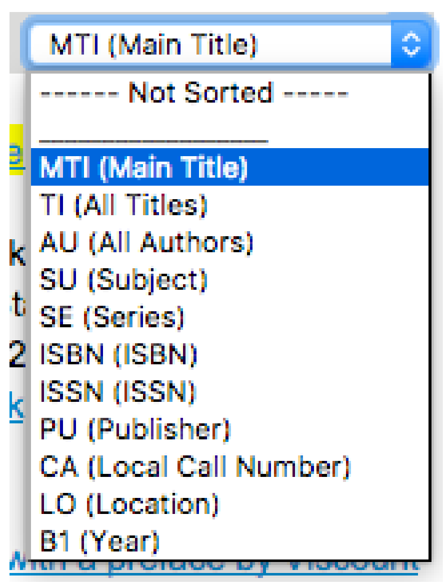


    - Portraits display – Title, creator, type, date, NGP number, and additional options (larger image, image zoom, buy a print, use this image).
    - Events display – Title, location, ticket price, description, keywords, and a buy tickets link (for those events that cost money).
    - Other pages – title and description (with the search term highlighted in the description).
- You can sort the search results (for portraits only) by "Gallery recommendation," "Date ascending," and "Date descending."
- The ranking of search results seems to be relevance (though this is not certain).

- There are some additional actions associated with the “Portraits” search. You can refine your search using a drop-down menu that lets the user select additional filters and you can conduct another search by using the search interface placed at the top of the results (you cannot save or select a subset of results).
- The search does not have a spell checker (“Earl of sandwhich” does not return any results, but the “earl of sandwich” returns 68 portraits and 7 people results) and it does not have stemming tools (a search for “paintings” only returns results for that term – not “paint,” “painter,” etc.). It does have a form of autocorrect/autosuggest – the search suggests past search terms when you start to type (it does not have phonetic tools or natural language processing tools).



- Query languages – it does support Boolean language, but it does not strip out “stop words.”
- The materials indexed for searching include full-text indexing and both navigation and destination pages (the “Other pages”), indexes events, by topic (metadata), and destination pages.
    - The search for earl also includes results that contain the world “early,” it indexes sections of a word.
- This system does have a type of vertical search – you can choose to see options in people, portraits, etc.



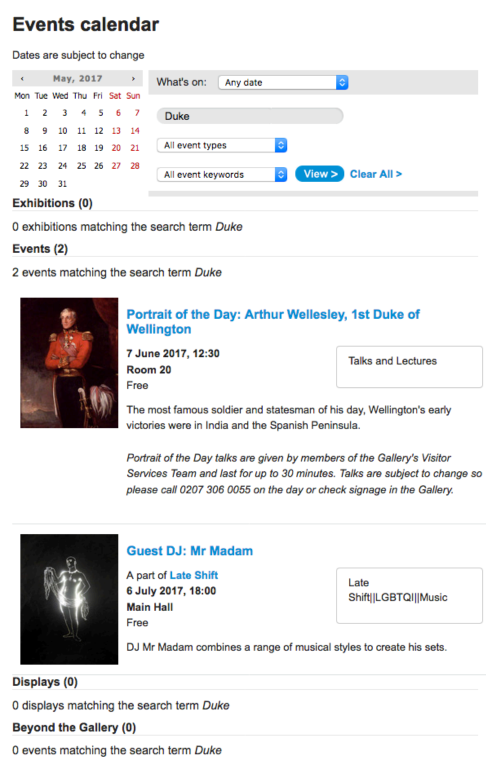
Road to 2012 search system:

- This is the least functional search system of them all.
- The search interface is a simple box located at the top of the page with a right arrow button.

- It does not have advanced search.
- The search results are organized into a short list (with just the title displayed). It does tell the user how many results were returned. All the titles are the same "the Commission" so you have no idea where to go.
- The search results are not sorted, ranked, there are no additional actions, there are no additional query builders or advanced search query languages.
- The content indexed is the description of the commissioned photos.

Archive and Library search system:

- This search system has three search types – simple, power, or filter
- Precision is more important for this search system. When you enter a search term you select the results that you would like to see. For example, if you search for "Horse" you can then select the results you want to see with the associated keywords (e.g., "Horse in art," "Horse breeds," etc.).
- The search interface (the main one) is located on the home page and above the search results. It has two sections, a drop-down menu with search fields and a search box. The search button is green with the word go on it. The simple search does not have the drop down menu option.



- The "Power Search" offers two search boxes and two corresponding drop-down menu (separated by Boolean search terms) and a green go button. The advanced search system also has additional options like search method, how many results to display, and limiting factors (see image above).
- The search results are displayed 20 per page (though that can be changed, see image above). The information displayed includes the title, publisher, locations, call number,

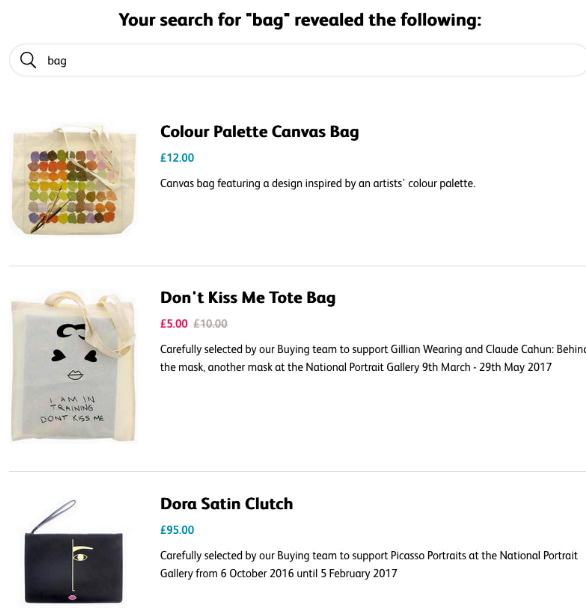
and related website. There are also images of the book/materials returned. There are also icons, which state what type of material is returned (book=book, eye=image, etc.).

- The search results are sorted by ascending order, but that can change to descending order. There is additional sorting option, by main title, all titles, all authors, subject, series, etc.
- There is no clear ranking system.
- There is a type of Best bets – the site has a list of "Special Searches," terms that are commonly searched and/or successful.
- There are many additional actions. Users can save a subset of the results to the "users list." They can save their search results (though not the search itself). You can print and email the search results. You can narrow the results down after it is first searched (you are taken to a list of terms related to the search term you have entered). A new search can be conducted using the search box at the top of the page that is prefilled with the original search term. If you return no results the search system offers tips for how to improve your query.
- It does not have spell checkers, phonetic tools, or natural language processing tools. It does have stemming features (e.g. a search for paintings returns results for paint). It does have a form of autocorrect/autosuggest – the search suggests past search terms when you start to type.
- Advanced query languages are supported in the search. Stop words are stripped out and Boolean language can be used.
- The content indexed for searching is the object metadata (title, publisher, location, call number, etc.). There is no full text indexing, etc.

Events Search system:

- The Events/Exhibition search has a keyword search.
- There are too little search results to confirm the algorithm and the recall and precision.

- The search interface is located at the top of the Events calendar page. It is a rounded search box with a "View >" button. There are additional filter options for the search (date, type, and keywords).

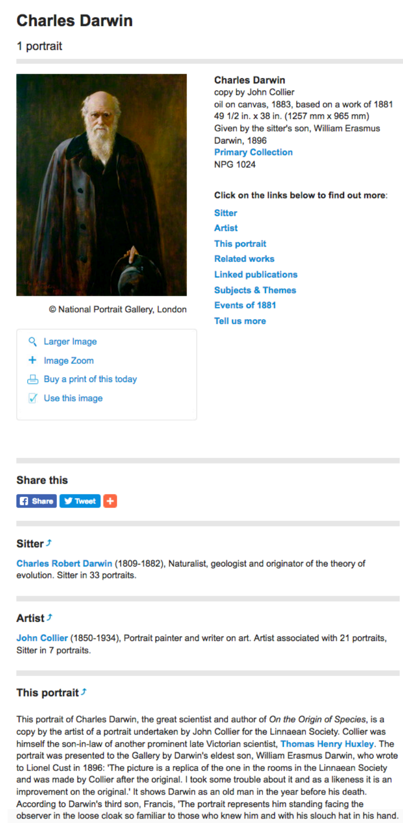


- There is no advanced search.
- The search results are displayed in a list form (organized under headings – Exhibitions, Events, Displays, and Beyond the Gallery). The information displayed includes the title, location, ticket price, description, keywords, and a buy tickets link (for those events that cost money).
- The search results are sorted by date (newest to oldest).
- You can revise you search using the search interface at the top of the event list. The search term that users have entered remains in the search box, so that the user can repeat/revise their search.
- There are not spell checkers, phonetic tools, natural language, etc.
- Advanced query languages are not supported.
- The content indexed for searching includes the metadata for the object (title, etc.), the keywords associated with the objects, the description of the object, etc.

The National Portrait Gallery store search system:

- The search system searches for object in the NPG store.
- The search interface is located at the top of the webpage (a circular search box with a magnify glass icon) and at the top of the search results.
- It does not have advanced search.
- The search results are displayed in a list form. There are images associated with each result, and the information includes the title of the object, the price, and a description.

- The sorting/ranking is unknown.
- The only additional action available is that the search interface is placed at the top of the results list (and populated with their search term).
- This search does not support spellcheckers or other query builders. For example a search for “Monarch” doesn’t return any results, but “monarchs” returns 15 results.
- Advanced query language is not supported
- The content indexed for searching is the item metadata (title, description, etc.). This is indexing by topic



Vocabulary systems

Metadata

Metadata is used throughout the site, used for both searching and browsing (see above).

Collection metadata:

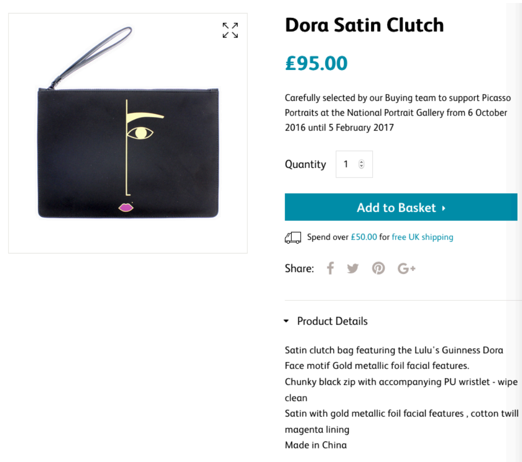
- The portraits contain a rich metadata field.
  - Descriptive metadata: Title, description, subject and themes, images, contextual information, and information about the sitter. Additionally, there are related resources about the object.
  - Structural metadata includes size, medium, etc.
  - Administrative metadata examples are artist, date, NPG call number, and provenance metadata.

Events, Exhibition, and Displays metadata:

- Administrative metadata – date, time and fee
- Descriptive metadata – title, description, and keywords.
- Structural metadata – room number

Object in the shop:

- Descriptive metadata – title, description, product details, images, etc.
- Structural metadata – medium (materials, etc.)
- Administrative metadata – location made, price, creator, etc.



Embedded metadata:

- The shop items have embedded metadata for better searching (in the <meta> field) – though it doesn't really say anything that is not availed on the main page (see image below). There was no descriptions/keywords in the <meta> field for events, portraits or people (due to time constraints not every item was examined).
- There were open graph embedded metadata – og:type and og:title

```
<meta name="description" content="Portraits of the British Monarchy from the National Portrait…split
Size: Approximately 370mm x 420mm (excluding handle) M">
```

Controlled vocabulary

The NPG uses a controlled vocabulary. There doesn't seem to be a synonym ring (a search for charcoal only returns portraits that have that word in it, not those that just have pencil – for example).

The A-Z indexes for sitters and artists act as authority files for the "correct" name structure. For example:

- Artists: "Sir Anthony van Dyck (1599-1641). 1018 Portraits" (listed under "V")
- Sitters: "Sir Frederick Augustus Abel, 1st Bt (1827-1902), Chemist and explosives expert. 8 Portraits"

The terms used in NPG were compared to numerous controlled vocabularies and thesauri to determine which standard it used (if any) – LCC, Art & Architecture (Getty), UNESCO, and HASSET.

Though the Archive and Library catalogue does use two standards - Art & Architecture (Getty) Authority file. For example, "Caravaggio, Michelangelo Merisi da (Italian painter, 1571-1610)" is formatted from the Art and Architecture Artist name Authority file. And had Library of Congress Subject Headings. For example, "Cardinals in art" or "Horse breeds."

Thesaurus

The subject and themes shows the NPG's thesaurus terms (see below).

- From the "Subjects and themes" sub page, a list of keywords that the portraits are tagged with have been collected and organized (through not all of the keywords). From this list you can decipher the thesaurus used for this digital museum.
- The excerpted thesaurus structure is organized below, it should be noted that these aren't all the keywords used to describe the portraits.
- You can see the semantic relationships between the terms. For example. You can see the hierarchical relationship within the "Pets and animals" category. Horses are nested under farm animals, which are nested under pets and animals. Related terms (associative relationships) have also been explained as well.
- It is important to note that this structure, while built off of a collection of keywords it was created. Some relationships can be identified through association within the images (e.g. "Film-shots and stage sets" almost always has the additional keyword "In character").

Accessory

- Art in art
- Books and libraries
- Carpet and textiles
- Clocks and timepieces
- Flowers and plants

Maps and globes
- Maps

Mirror and reflections

Pets and animals
- Pets and animals – Birds
- Pets and animals – Cats
- Pets and animals – Dogs
- Pets and animals – Farm animals
  - Pets and animals – Horses
- Pets and animals – Fish and sea life
- Pets and animals – Mythical Beasts
- Pets and animals – Rabbits and rodents
- Pets and animals – Reptiles and amphibians
- Pets and animals – Wild and exotic animals

Activity
- Dancing
- Drinking and eating (written as "Eating and drinking" in portrait metadata)
- Making art RT Art in art
- Making music
- Reading RT Books and libraries
- Sleeping
- Smoking
- Writing

Dress
- Eyeglasses and spectacles
- Fans
- Gloves and gauntlets
- Hats and head attire
- Jewellery

Jewellery – Amulets and religious symbology

Jewellery – Bangles and bracelets

Jewellery – Broaches, buttons, and buckles

Jewellery – Crowns and tiaras

Jewellery – Hair accessories

Jewellery – Earrings

Jewellery – Livery chains and badges

Jewellery – Pearls

Jewellery – Pendants and necklaces

Jewellery – Precious stones

Jewellery – Rings

Masks and disguises

Umbrellas and parasols

Genre

Children

Couples

Double portraits

Family

Family portraits

Group portraits

Nudes and naked figures

Body

Royal babies BT Children

Self-portrait

Weddings

Wedding inspiration

Themes

Artists and their studios RT Art in art

Buildings and architecture

Stairways and walkways

Diversity

Events and occasions

Film-shots and stage sets

In character

Gardens

Flowers and plants

Snow

Words and inscriptions

These terms do not follow standards – Art and Architecture (Getty), Library of Congress Subject Heading, HASSET, etc.

Faceted Classification

The blog has a type of faceted classification. Users can limit the blog posts they want to see according to the tags that have been applied to them (see image). The advanced search feature also has facets that can narrow down the search results.

The National Portrait Gallery Sources

Portrait Gallery Corporate Plan 2016-2019 (2016). Our ten year vision. Retrieved from: http://www.npg.org.uk/assets/files/pdf/corporate/busplan20162019.pdf

Digital Strategy (Nov. 2016). Digital Aims. Retrieved from: http://www.npg.org.uk/assets/files/pdf/strategic-plan/NPG_Digital_Strategy_Digest_v4_1.pdf

I would like to note that all images from the National Portrait Gallery website are only being used for educational purposes.