Artificial Intelligence Literacy: A Proposed Faceted Taxonomy

Ali Shiri, School of Library and Information Studies, University of Alberta, Canada

Purpose: The purpose of this paper is to propose a taxonomy of Artificial Intelligence literacy to support AI literacy education and research.

Design/methodology/approach: This study makes use of the facet analysis technique and draws upon various sources of data and information to develop a taxonomy of AI literacy. The research consists of the following key steps: a comprehensive review of the literature published on AI literacy research, an examination of well-known AI classification schemes and taxonomies, a review of prior research on data/information/digital literacy research, and a qualitative and quantitative analysis of 1,031 metadata records on AI literacy publications. The KH Coder 3 software application was used to analyze metadata records from the Scopus multidisciplinary database.

Findings: A new taxonomy of AI literacy is proposed with 13 high-level facets and a list of specific subjects for each facet.

Originality: The proposed AI literacy taxonomy offers a new and original conceptual framework that builds on a variety of different sources of data and integrates literature from various disciplines, including computing, information science, education, and literacy research.

Research limitations/implications: The proposed taxonomy may serve as a conceptual AI literacy framework to support the critical understanding, use, application, and examination of AI-enhanced tools and technologies in various educational and organizational contexts.

Practical implications: The proposed taxonomy provides a knowledge organization and knowledge mapping structure to support curriculum development and the organization of digital information.

Social implications: The proposed taxonomy provides a cross-disciplinary perspective of AI literacy. It can be used, adapted, modified, or enhanced to accommodate education and learning opportunities and curricula in different domains, disciplines, and subject areas.

Keywords: Artificial intelligence literacy, AI literacy, Taxonomies, AI literacy taxonomy, AI literacy conceptualization, Artificial intelligence

1. Introduction

With the rapid development and popularity of generative AI tools, applications, and search engines particular attention has been drawn to the critical importance of AI literacy and its integration into a variety of educational and training opportunities for K-12 and higher education. A number of recent scoping and literature review studies have examined the growth of publications and research on AI literacy and its implications for various disciplines, subject areas, and educational contexts. In a study of artificial intelligence literacy in adult and higher education, Laupichler et al. (2022) highlight a gap in the literature of AI and stress the need for a current and cross-disciplinary AI basic vocabulary that should be mastered by every person regardless of their profession or discipline. This paper aims to address this

vocabulary gap in the literature by proposing a taxonomy of AI literacy using the facet analysis technique. The overarching goal of proposing such a taxonomy is to provide a conceptual framework for AI literacy knowledge, competencies, and skills in order to support AI literacy education and research. Its development is rationalized by the fact that high level subject categories and terms provide a conceptual map of the AI literacy domain and facilitate learning, curriculum development, as well as the identification of the gaps in educational contexts and research topics. The proposed taxonomy is aimed to serve as a conceptual framework and draws upon the existing literature on AI literacy in various contexts, including in K-12, adult education, and higher education. This conceptual framework does not intend to provide applications for specific educational contexts such as K-12 or higher education. Rather, the goal is to provide a high level conceptual framework for understanding and communicating the key aspects and facets of AI literacy. The reason for this approach lies in the fact that many of the research and review publications that have proposed AI literacy models and frameworks have made use of a combination of AI and literacy topics sometimes navigating multiple educational contexts and overlap in their treatment of topics as well as knowledge, competencies, and skills. The proposed taxonomy in this paper has the potential to be further expanded and customized for various user communities and educational contexts.

2. Prior Research and Context

2.1 Taxonomies of Digital literacy, Information Literacy, Data Literacy

In a theoretically and conceptually well-laid-out review of various types of literacy, Stordy (2015) notes that the digital technologies and developments have transformed our conceptualization of literacy and what it means to be a literate person. He highlights 35 different types of literacy from computer literacy, information literacy and digital literacy, to newer conceptions like transliteracy, metaliteracy and multimodal literacy. Drawing upon the work of Lankshear and Knobel (2007) and Street (2007), he proposed a new taxonomy of literacies that reflect the recent conceptualizations and variety of literacy in relation to digital technologies. He argues that while these categories are useful to conceptualize different types of literacy, they are overlapping and are not mutually exclusive categories. Stordy's facets (2015) include:

> Autonomous-conventional perspectives, including Internet and the information explosion, information literacy falls under this category.
>
> Autonomous-Peripheral perspective: signifies literacy related to digital technologies, digital literacy falls under this category
>
> Autonomous- Paradigm perspective: implying literacies that support new ways of being literate, including information literacy 2.0, metaliteracy and programming literacy
>
> Ideological-Conventional perspective: conceptualizes literacy in the context of social practices and not directly related to digital technologies, early sociocultural perspectives of information literacy fall within this category.
>
> Ideological-Peripheral perspective: refers to literacies that cover both sociocultural and digital technology perspectives. Internet literacy has been used as an example of this type of literacy.
>
> Ideological-Paradigm perspective: conceptualizes the type of literacy that has stemmed from new social practices as a result of the emergence of new technologies. Digital literacies and transliteracy could be categorized under this perspective.

One may argue that due to the complex and evolving nature AI technologies and what new learners need to learn and understand, the conceptions of AI literacy may benefit from the above categories, and more specifically from the concept of transliteracy and ideological-peripheral perspective.

As the pressing need for the development of a set of AI literacy competencies and skills emerges, research and development activities are underway to draw upon prior research on various types of literacy, including information literacy, digital literacy, internet literacy, network literacy, data literacy, computer literacy, algorithmic literacy, and news and social media literacy. Bawden provides a comprehensive discussion and definitions of different types of literacy and how they overlap or complement one another (Bawden, 2021). One of the widely used framework for the development of these literacy competencies is Bloom's taxonomy with its key categories, namely Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation (Krathwohl, 2002). Variations, adaptations, and modifications of these key categories of literacy have been used to develop taxonomies for information literacy competencies (Vitolo & Coulston, 2002; Keene et al., 2010), digital literacy competencies (Riel et al., 2012; Reddy et al., 2023), data literacy competencies (Calzada & Marzal, 2013; Erwin, 2015; Schuff, 2018), and algorithmic literacy (Ridley & Pawlick-Potts, 2021).

2.2 AI and AI Literacy

2.2.1 AI definitions

Before we review prior research and development in the area of AI literacy, it is important to provide a brief discussion of the definition for the term Artificial Intelligence (AI). There have been numerous definitions and conceptualization of AI proposed since the term was coined. In this section, we briefly review a few definitions to set the stage for the discussion of AI literacy. Historically speaking, the field that we currently know as artificial intelligence is attributed to the work of visionary researchers such as McCarthy, Minsky, Rochester and Shannon, and Simon who participated in Dartmouth meeting to coin the term artificial intelligence in 1955 (McCarthy et al., 1955). Wikipedia (artificial intelligence n.d.) defines AI as "the intelligence of machines or software, as opposed to the intelligence of living beings, primarily of humans. It is a field of study in computer science that develops and studies intelligent machines. Such machines may be called AIs." McCarthy (2007) defines AI as "It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable." In a detailed and critical review of the definitions for the term artificial intelligence, Wang (2019) offers a definition of AI focusing on the conceptualization of intelligence as follows: "Intelligence is the capacity of an information-processing system to adapt to its environment while operating with insufficient knowledge and resources." One of the most comprehensive definition of AI is attributed to the European Commission High-Level Expert Group on Artificial Intelligence as follows:

> "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans (2) that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions."(Samoili et al., 2020)

There is also growing literature on competing vision around artificial general intelligence (AGI) or human like intelligence and narrow AI or weak AI that are able to accomplish specific tasks. For a critical and conceptual review of these two types of AI, readers are referred to Korteling et al. (2021).

2.2.2 AI literacy

Long and Magerko (2020) provide a concrete definition of AI literacy and propose a set of core AI competencies based on an examination of interdisciplinary literature. They define AI literacy as "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace." Their conceptual framework consists of five key questions, namely: What is AI?; what can AI do?; How does AI work?; How should AI be used?; and How do people perceive AI?" Based on these questions they propose a list of 16 AI competencies as listed in Table 1.

Ng et al. (2021a & 2021b) carried out an exploratory review of literature on the definitions and conceptualizations of AI literacy proposed by researchers. Using Bloom's taxonomy, they proposed an AI literacy framework with six key facets, namely Know, Understand, Apply, Analyze, Evaluate, Create (Table 1). They note that the most common approach to AI literacy is to base the conceptualization on other types of literacy. Their important observation from the analysis of literature is that rather than focusing AI literacy solely on how to use applications, learners should be taught the underlying AI concepts for their future career and ethical concerns in the adoption and use of AI applications to be informed and responsible citizen. In a recent study, Rizvi et al. (2023) conducted a systematic review of literature to examine research with empirical evidence outcomes for teaching and learning AI in K-12 between 2019 and 2022. They provided a conceptual analysis of the topics discussed in the literature and proposed the AI Learning Levels Framework consisting of four high level categories, namely Engines, Models, Applications, Social and Ethical Considerations. Table 1 provides a tabular overview of the work on competencies, frameworks, and high level categories in AI literacy proposed Long & Magerko (2020), Ng et al., 2021a & 2021b, and Risvi et al., 2023).

Long & Magerko (2020)	Ng et al. (2021a & 2021b)	Rizvi et al. (2023)
-Competency 1: Recognizing AI -Competency 2: Understanding Intelligence -Competency 3: Interdisciplinarity -Competency 4: General vs. Narrow -Competency 5: AI's Strengths & Weaknesses -Competency 6: Imagine Future AI -Competency 7: Representations -Competency 8: Decision-Making -Competency 9: ML Steps -Competency 10: Human Role in AI - -Competency 11: Data Literacy -Competency 12: Learning from Data -Competency 13: Critically Interpreting Data) -Competency 14: Action & Reaction -Competency 15: Sensors -Competency 16: Ethics	-Know: Use information in new situations (Copy reproduce, recall and memorize AI concepts -Understand: Explain ideas or concepts (describe, explain, interpret and demonstrate the meaning of AI -Apply: Use information in new situations (Execute, implement, use and apply AI applications in different contexts -Analyze: Draw connections among ideas (Organize, compare, decompose, an abstract an AI problem -Evaluate: Justify a stand or decision (Appraise, predict, detect and justify decisions with AI applications -Create: Produce new or original work (Design, assemble, construct, build and develop AI applications)	-Social, ethical considerations level: e.g. knows about the idea of bias in machine learning (ML), understands that artificial intelligence is not magic and that machines are not self-deterministic. -Applications level: e.g. knows that how systems that include AI components, can design an application that includes ML image recognition -Models level: e.g. can explore an ML model that was created by someone else, understands the process of selecting and cleaning data needed to train a simple ML model -Engine level: e.g. can explain how a decision tree can be used to classify items, can explain in simple terms how a neuron works with relationship to learning about ML

Table 1. AI literacy competencies and frameworks (table by author)

Using a quantitative assessment approach, Wang et al. (2022) developed a quantitative scale for obtaining accurate data regarding the AI literacy of ordinary users. They made use of four key constructs to conduct their assessment, namely Awareness, Usage, Evaluation, Ethics. They developed a set of 12 measurement items focusing on users' knowledge of AI technologies and applications, their learning to use and evaluate AI tools, AI use in problem solving, and AI ethical and privacy issues. Laupichler et al. (2023) conducted a more comprehensive Delphi study to develop an assessment tool for non-experts' AI literacy with 38 items. The key assessment questions included such terms as AI technologies, human and AI, general and narrow AI, AI applications in media and games, AI weaknesses and strengths, AI risk and advantages, knowledge representation, machine learning, machine learning models, supervised and unsupervised learning, deep learning, artificial neural network, critical evaluation of AI applications, the role of data, AI assisted sensor data gathering, computer vision, natural language processing, AI ethical and legal issues, AI biases, AI impact on society, history and definition of AI, explainable AI, big data, what is an algorithm, and its recent increased popularity. Other researchers have proposed integrated data and AI literacy frameworks that include such literacy constructs as knowledge, skills, and attitudes and that inform the development of curricula and educational standards (Schüller, 2022). Kong et al. (2021) reported an evaluation of an AI literacy course for university students with diverse study backgrounds consisting of four major concepts, namely artificial intelligence, machine learning, supervised learning and unsupervised learning. They argued that AI literacy should include three main components: AI concepts for understanding and developing AI literacy, using AI concepts for evaluation in order to evaluate AI technology, and using AI concepts for understanding the real world through problem solving at home and in the workplace. Cardon et al. (2023) propose a set of guiding questions for the development of AI literacy for business communication and writing, focusing on such concepts as Application (understand AI tools), Accountability (take responsibility for content reliability), Authenticity (focus on genuine communication), and Agency (retain decision-making right). Through a systematic literature review of AI literacy in K-12, Casal-Otero et al. (2023) develop a taxonomy of approaches to AI learning in K-12 including,

- learning experiences focused on understanding AI: Learning to recognize artefacts using AI, learning about how AI works, learning tools for AI, learning life with AI
- proposal for implementation of AI learning in K-12: AI literacy curriculum design, AI as a subject in K-12 education, student perspective on AI literacy, teacher training in AI, AI literacy support resources, gender diversity in AI literacy

Stolpe, K., & Hallström (2024) review various AI literacy frameworks and propose new framework for AI literacy for technology education. Their conceptual framework consists of a matrix of technological scientific knowledge, technical skills, and socio-ethical technical understanding. They argue that while technical skills such as programming competencies appear in the reviewed AI literacy frameworks, they are less emphasised.

2.2.3 AI Taxonomies

In proposing a taxonomy of AI literacy, it is also important to consider other studies that have focused on the development of a taxonomy for the domain of AI in general. This is particularly relevant to see if any knowledge structures or taxonomies have incorporated AI literacy or aspects of AI literacy in their terminology. In this section, we present a brief review of the previous research and development in this area. One of the most comprehensive taxonomies of AI is represented in the Association for Computing Machinery (ACM) Classification (Table 2). While this classification provides a detailed list of subject terms and topics for the classification of AI publications and documents, its treatment of AI literacy is very limited. One of the key facets of this classification is Social and Professional Topics, under which a

number of AI literacy related themes and terms may be found, including such terms as *computing literacy, computational thinking, computing education, social aspects of security and privacy, human and societal aspects of security and privacy*, as well as the term *privacy policies* under *computing and technology policy*.

Buchanan et al. (2013) developed one of the first comprehensive taxonomies of AI concepts based on 130 influential publications to organize AI research and publications in a virtual archive. Their taxonomy covers broad topics on Education, and Ethics and Social Issues, with Education which focuses specifically on the application of AI in education with such topics as *automated grading, instructional theory,* and *intelligent tutoring* and the category of *ethics and social issues* covering the terms *robot ethics, security and privacy and social issues*. In a bibliometric study of the growth of publications on AI, Niu et al. (2016) conducted a keyword analysis of AI research publications from 1990-2014 and proposed a number of frequently used terms in the publications on AI. They further analyzed the longitudinal trends of keywords occurring in the AI publications for the time period. Their term frequency analysis has a particular focus and emphasis on technical aspects, including methods, models, and applications as the central components of AI research.

In addition to AI literacy studies, Martınez-Plumed et al. (2018) developed 9 facets for the domain of artificial intelligence to provide a systematic and analytical framework for the shape and boundaries of AI. The key facets are listed in Table 2. In their framework, *functionality facet* focuses on the functionality of AI systems such as knowledge representation, reasoning, learning, communication, perception, and action. The *referent facet* makes distinction between definitions of AI from the anthropocentric or universal (theoretical) perspectives. In other words, the human edge denotes human like AI that focuses on automating the human processes. The universal aspect aims to capture theoretical and abstract processes. Examples of terms under facets number 3 to 9 are briefly listed in Table 2.

Table 2 provides a tabular overview of the AI classification schemes, taxonomies, facets proposed by Buchanan et al. (2013), Niu et al. (2016), Martinez-Plumed et al. (2018) and the ACM Classification.

Buchanan et al. (2013)	**Niu et al. (2016)**	**Martınez-Plumed et al. (2018)**	**ACM Classification**
- AI Overview - Applications - Cognitive Science - Education - Ethics and Social Issues - Games and Puzzles - History - Interfaces - Machine Learning - Natural Language - Philosophy - Representation and Reasoning - Robots - Science Fiction - Speech - Systems and languages - Vision - Web and AI	- Artificial Intelligence - Artificial Neural network - Genetic algorithm(GA) - Expert System (ES) - Optimization - Prediction - Classification - Design - Fuzzy Logic - Multi-agent system - Simulation - Support Vector Machine - Machine Learning(ML) Diagnosis	- Functionality facet (with edges ‘techniques’, ‘applications’, ‘tasks) - Referent facet (with edges ‘human’ and ‘universal’) - Generality facet (narrow vs. general AI) - Location facet (centralized or distributed decision making by an agent) - Embodiment facet (physical or virtual)	Artificial intelligence Natural language processing - Information extraction - Machine translation - Discourse, dialogue and pragmatics - Natural language generation - Speech recognition - Lexical semantics - Phonology / morphology - Language resources Knowledge representation and reasoning - Description logics Semantic networks - Nonmonotonic, default reasoning and belief revision - Probabilistic reasoning - Vagueness and fuzzy logic - Causal reasoning and diagnostics - Temporal reasoning - Cognitive robotics - Ontology engineering - Logic programming and answer set programming

	- Data mining - Swarm intelligence(SI) - Particle swarm optimization - Decision support system(DSS) - Pattern recognition - Knowledge-based system(- Case-based reasoning(CBR) - Distributed artificial Intelligence(DAI) - Knowledge representation(KR) - Management - Identification - Decision making - Fault Diagnosis - Computational intelligence - Recognition - Ontology	- Paradigm facet (discrete, continuous, probabilistic) - Actor facet (academia, industry, government, independent) - Character facet (empirical or theoretical) - Nature facet (disciplines: technology, engineering, science, philosophy)	- Spatial and physical reasoning - Reasoning about belief and knowledge Planning and scheduling - Planning for deterministic actions - Planning under uncertainty - Multi-agent planning - Planning with abstraction and generalization - Robotic planning - Evolutionary robotics Search methodologies - Heuristic function construction - Discrete space search - Continuous space search - Randomized search - Game tree search - Abstraction and micro-operators - Search with partial observations Control methods - Robotic planning - Evolutionary robotics - Computational control theory - Motion path planning Philosophical/theoretical foundations of artificial intelligence - Cognitive science - Theory of mind Distributed artificial intelligence - Multi-agent systems - Intelligent agents - Mobile agents - Cooperation and coordination Computer vision - Computer vision tasks - Biometrics - Scene understanding - Activity recognition and understanding - Video summarization - Visual content-based indexing and retrieval - Visual inspection - Vision for robotics - Scene anomaly detection Image and video acquisition - Camera calibration - Epipolar geometry - Computational photography - Hyperspectral imaging - Motion capture - 3D imaging - Active vision Computer vision representations - Image representations - Shape representations - Appearance and texture representations - Hierarchical representations Computer vision problems - Interest point and salient region detections - Image segmentation - Video segmentation - Shape inference

			- Object detection - Object recognition - Object identification - Tracking - Reconstruction - Matching Machine learning - Learning paradigms - Supervised learning - Unsupervised learning - Reinforcement learning - Multi-task learning - Learning settings - Machine learning approaches - Machine learning algorithms Social and Professional Topics - Computing education - Computational thinking - Computing and technology policy - Privacy policies Security and Privacy - Human and Societal Aspects of Privacy and Security - Privacy protections

Table 2. AI classification schemes, taxonomies, facets (table by author)

In a recent bibliometric study of publications on AI literacy to examine the main characteristics of the current scientific output, Tenório et al. (2023) note that since 2018 there has been an increase in the number of publications on AI literacy. Their keywords co-occurrence analysis shows that artificial intelligence and machine learning are the most frequently occurring terms. The key term clusters that were created in their analysis included such topics *as big data, data, deep learning, and ethics*. Other key term clusters included *education, data literacy, AI education, AI literacy, and machine learning education.* While their study provides a broad bibliometric overview of the publications on AI literacy, there is little terminological or taxonomic information about the specific themes, terms, and keywords on different aspects of AI literacy.

In addition to the above taxonomies of AI, there are a number of other knowledge structures and knowledge organization systems that have been used to organize AI publications and research. One example is the list of subject terms assigned to AI publications in the Academic Search Complete, a multi-disciplinary database of scholarly publications. A review of the AI related subject terms in the Academic Search Complete subject term list shows that most of the terms overlap with the terms that are listed in Table 2. There are only a couple of specific terms that fall under the broad theme of AI literacy, namely AI and ethics and AI in education.

While above taxonomies provide a comprehensive overview of the subject categories and terms related to AI, their treatment of AI literacy as an emerging area is limited. For instance, Buchanan et al. (2013) contains terms such as education and ethics and social issues, but does not provide a high degree of specificity for various aspects of AI literacy education. Although Martinez-Plumed et al. (2018) briefly address human and policy aspects of AI, their focus is not on educational and literacy aspects of AI and as a result terms related to human's learning about and understanding of AI tools and functions are not included. While the ACM Classification provides a comprehensive and detailed coverage of AI related topics and terms, it lacks specificity in treating terms and subject categories related to AI literacy and education. For instance, it contains terms such as social and professional topics such as computational

thinking, computing education, social and human aspects of privacy and security, but does not cover specific aspects of AI education in various contexts and in relation to various literacy frameworks such as digital and information literacy. In this paper, the goal is to propose a taxonomy that covers educational and literacy facets of AI as well as its social, ethical, and technological aspects.

3. Methods

There are different approaches, methods, and subject term sources reported in the literature for creating a taxonomy or controlled vocabulary for a subject area. Depending on the evolutionary or stable nature of a subject domain, a wide range of sources of subject categories, themes, and terms could be used in building a taxonomy. These sources may include existing controlled vocabularies, taxonomies, subject classification schemes; published literature on the subject; domain-specific databases, websites, digital libraries, and online materials; users' search terms in various search engines and content management systems; expert opinions (Broughton, 2006; Zaharee, 2013); and ; bibliometric reports. Depending on the scope and multidisciplinary nature of the subject domain, taxonomy developers may use one or a combination of these sources for taxonomy creation. In this study, we used a combination of sources of subject categories and terms. In order to provide a holistic approach in identifying existing AI taxonomies and knowledge structures that may contain topics related to AI literacy, a two-pronged method was adopted to maximize the findability of existing controlled vocabularies.

The first was to identify existing controlled vocabularies in the broad area of AI. The author was already familiar with the ACM Classification as he had used it in previous research. Furthermore various scholarly databases were examined to ascertain if they have controlled vocabularies that they use for indexing. These included, Academic Search Complete, Scopus multidisciplinary database, and Web of Science databases. While these databases cover and index AI publications, their controlled vocabularies are not as comprehensive and domain-focused as the ACM Classification.

The second approach was to conduct literature searches in Scopus and ACM Digital Library in order to identify controlled vocabularies, taxonomies, term lists, and thesauri in the area of AI and AI literacy. These searches contributed both to the identification of existing taxonomies and controlled vocabularies as well as articles and systematic reviews that addressed AI literacy conceptualizations, models, frameworks, and taxonomies. These searches also allowed us to identify bibliometric studies that had been conducted on AI literacy research and publications.

Table 3 shows the sources of subject terms and themes and sources used.

Sources of subject terms and themes	**Sources used**
Existing classification schemes, taxonomies, and proposed subject categories	ACM Digital Library Scopus multidisciplinary database Web of Science databases Academic Search Complete ACM Classification Buchanan's taxonomy of AI (2013) Facets of artificial intelligence (Martınez-Plumed et al. (2018) Bibliometrics-based reports (Niu et al., 2016; Tenório et al. (2023)
AI literacy skills and competencies reported in the literature	Long & Magerko (2020) Ng et al. (2021a & 2021b)

	Rizvi et al. (2023)
Scopus metadata records on AI and literacy	1,031 metadata records on AI and literacy

Table 3. Sources of subject terms and themes used in this study (table by author)

In addition to the above key sources, the content of several recent scholarly papers on AI literacy was examined and analyzed to gain a deeper insight into the emerging themes and topics. These resources are all cited in this paper.

Scopus data and search strategy

In addition to the published literature, an analysis of metadata records of 1,031 publications on AI and literacy was conducted. The Scope multidisciplinary database was utilized to conduct searches on AI and literacy. The search strategy focused specifically on the search statement AI OR Artificial Intelligence and Literacy in the Scopus fields of 'Article Title, Abstracts, keywords' rather than expanding the search with synonymous or conceptually related terms. All the searches were conducted on July 25, 2023. The detailed search strategy is presented in the box below.

```
( TITLE-ABS-KEY ( "artificial intelligence" ) OR TITLE-ABS-KEY ( ai ) AND TITLE-ABS-
KEY ( literacy ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) ) AND ( LIMIT-
TO ( DOCTYPE , "ar" ) OR LIMIT-TO ( DOCTYPE , "cp" ) OR LIMIT-
TO ( DOCTYPE , "re" ) OR LIMIT-TO ( DOCTYPE , "ch" ) OR LIMIT-
TO ( DOCTYPE , "cr" ) OR LIMIT-TO ( DOCTYPE , "ed" ) OR LIMIT-TO ( DOCTYPE , "bk" ) )
```

In total, 1,099 records were found that met the search terms AI or Artificial Intelligence AND Literacy. The search was narrowed down to the English language and also the Scopus document types such as 'letter', 'notes', 'retracted' etc. were excluded from the records. In total, 1,031 records were retrieved and stored.

Text analysis

The KH Coder 3 text analysis tool was employed to analyze the metadata records and develop term and theme clusters. KH Coder is a free software application for quantitative content analysis and text mining. It has also been utilized for computational linguistics (https://khcoder.net/en/). KH Coder 3 provides a number of text analysis tools and techniques, including term frequency, co-occurrence analysis, cluster analysis, multidimensional scaling, and correspondence analysis methods. The Word Cluster for term extraction feature was used to obtain a list of word clusters. Using the Word Cluster feature, a list of over 21,000 key phrases were extracted. Since these key phrases are ranked in order of frequency, only the top 5,000 phrases were manually examined to inform the development of the proposed taxonomy. This decision was rationalized by the fact that the top 5,000 phrases had a frequency range sufficient to identify frequently recurring themes and topics. This decision was also rationalized by the fact that since the article titles, abstracts, and keywords were used in this analysis, additional review and analysis of terms beyond the 5,000 phrases would have resulted in redundant phrases. As a result, we focused on the recurring themes and topics that would fall under the most frequently used terms.

The top 100 frequently used phrases in the abstracts of the analyzed Scopus metadata records are listed in Table 4.

artificial intelligence	ai curriculum	training data

ai literacy	artificial intelligence literacy	internet literacy
digital literacy	digital media	medical information
health literacy	digital transformation	health systems
information literacy	computer science	understanding of ai
ai education	case study	medical education
data literacy	technology literacy	data management
machine learning	fake news	digital health technologies
health care	artificial intelligence education	digital age
big data	literacy learning	public health
use of ai	development of ai	scientific literacy
digital health	new technologies	literacy level
artificial intelligence technology	deep learning	data design
information technology	ai use	digital health education
health information	media literacy education	language learning
media literacy	data analysis	field of education
ai systems	digital literacy skills	application of ai
social media	engineering education	literacy development
digital technology	smart education	advancement of artificial intelligence
financial literacy	health care education	basic literacy
digital technologies	data science	teaching process
ai technology	ai applications	digital data
ai literacy education	digital health literacy	ai perception
ai knowledge	information literacy of teachers	literacy skills
digital education	media education	k-12 education
ai learning	digital economy	online learning
ai technologies	role of ai	artificial intelligence technologies
digital skills	students ' information literacy	knowledge of ai
future research	digital learning	public ai literacy
literacy education	information literacy education	medical students
technological literacy	university students	digital health technology
students ' ai literacy	digital literacy education	active learning
computer literacy	digital media literacy	health care professionals
		quality of care

Table 4. Top 100 frequently used phrases in the Scopus metadata records (table by author)

A review of the list of top frequently used terms in the AI and literacy literature (Table 4) shows that researchers frequently reference different types of literacy associated with information, data, digital environment, media, internet, and technology as they discuss and conceptualize the various aspects and facets of AI literacy. This approach to the emerging discussion of AI literacy is in line with Stordy's argument (2015) that the definition of different types of literacy related to the digital environment could benefit from concept of transliteracy and digital literacies.

4. Analytical Framework

The general principles of facet analysis was used to develop a set of high level categories as well as specific themes and subjects that represent specific aspects of AI literacy. Facet analysis was introduced by Ranganathan (1962) as a model for the development of knowledge organization systems such as library classifications and thesauri. Based on this theory, Aitchison et al. (2002) provide a more specific and descriptive set of fundamental categories that are useful as a practical basis for facet analysis. These are as follows:

1. Entities, things, and objects subdivided by characteristics and function
2. Actions and activities
3. Space, place, location, and environment
4. Time
5. Kinds or types; systems and assemblies; applications and purposes

La Barre, K. (2010) provides a detailed review of the facet analysis technique and elaborates on its process that entails the definition of the subject field and facet formulation. Variations of the facet analysis technique have been used to develop a conceptual framework for making sense of big data (Shiri, 2014), video game genres (Lee et al., 2014), software reuse (Albrechtsen, 2015), and research data management and data curation (Sandy et al., 2019). As mentioned in the methods section, in order to develop a set of facets to categorize topics and themes on AI literacy, the following sources were used:

- the topics and themes related to AI literacy in AI taxonomies
- AI literacy publications
- Scopus metadata records on AI and literacy

In the process of facet analysis and facet formulation using the above fundamental categories, two different conceptual frameworks were considered. In a comprehensive review of information and digital literacy definitions, Bawden (2001) concludes that while the skills and competencies can be used in defining information and digital literacies, understanding, meaning and context should be central to the way in which we define information and digital literacy. This holds true for developing a conceptual framework and a taxonomic structure for AI literacy that can encompass core principles, skills, and practices without confining it to certain technologies or skill-based literacies. The second framework, proposed by Oeldorf-Hirsch & Neubaum (2021), provides a conceptual framework using the cognitive, affective, and behavioural dimensions that defines algorithmic literacy as follows:

> *Cognitive*: knowing, understanding, awareness, knowledge (cognitive comprehension of their existence and functioning)
> *Affective*: feeling, sensing, aversion, appreciation (thoughts, attitudes)
> *Behavioural*: Doing, using, engaging, skills

In addition to the two conceptual frameworks referenced above, the analysis of the abstracts of the publications metadata records in this study also showed that there were frequently used terms such as knowledge, concepts, conceptual understanding, attitudes and perceptions, skills, and applications. The term frequency analysis in this study further confirmed the importance of creating high level facets and subject categories that could provide a general concept map with the potential to encompass a more nuanced and detailed account of the terms associated with AI literacy. It should be noted that the proposed taxonomy presented in the next section is the result of the facet analysis of the terms found in the existing controlled vocabularies listed in Table 3, frequently used terms in the metadata records from the Scopus database listed in Table 4 and the high level facets discussed in this section. The proposed taxonomy is a new conceptualization of AI literacy that while it draws upon and makes use of the above mentioned sources, the conceptualization is not an exact extraction or adaption of these sources. Rather, it represents an abstract and facet-based representation of the key themes, topics, and terms on AI literacy that the existing controlled vocabularies may cover only partially and not with a particular focus on AI literacy as an emerging construct.

5. A Proposed Taxonomy of AI Literacy

The result of the facet analysis of AI literacy terms, themes, concepts found in the previous publications and Scopus metadata records on AI and literacy is presented in Table 5. As can be seen, 13 high level facets for AI literacy are proposed based on the facet analysis fundamental categories listed in the analytical framework section, namely activities and operations, entities, individuals and organizations, objects, environments (physical and digital), time, kinds and types. Using the examined literature and metadata records, a list of subjects was created for each facet to provide a more detailed and delineated list of conceptually related subjects, concepts, and ideas.

Facet	**Subjects**
Conceptualization of AI (understanding, awareness, knowledge)	AI Concepts, definitions, and conceptual frameworks, current popularity, attitudes, motivation and perceptions towards AI, digital economy
Contextualization of AI Literacy	Technology literacy, digital literacy, data literacy, information literacy, internet literacy, social media literacy, and algorithmic literacy, integrated frameworks that use data/information/digital literacies as a foundation, AI literacy curriculum and learning applications in various educational and pedagogical settings (K-12, colleges and universities)
Applied knowledge	Machine learning, supervised and supervised learning, deep learning, natural language processing, neural networks, knowledge representation and reasoning
AI applications and its interdisciplinarity	Applications in various disciplines, domains, and industries Applications in business, health, education, media, gaming, music etc., Applications in audio, video, speech, vision, text, object, and image analysis, recognition, and classification
Skills: Use, recognize, interact and engage with, critically examine AI tools	Technologies that use AI, pros and cons of AI-based systems (including safety, reliability, and trustworthiness), distinction between general and narrow AI, impacts of AI, decision support and problem solving using AI technology, physical and digital tools and technologies that use AI
Data	Emphasis on data as a key component of AI tools and machine learning applications, big data, data in different contexts such as in search engines, businesses, banking, large language models (ChatGPT, Google Gemini, Perplexity.ai), robots, etc.
Ethical, legal, social aspects, issues, concerns	Data privacy and protection, security, confidentiality, transparency, accountability, explainability, algorithmic bias, discrimination, human intelligence, human-centred AI, Human-in-the-loop, autonomy, agency, copyright, intellectual property, generative authorship, misinformation and information integrity and credibility
Environment (physical)	Libraries, archives, museums, publishers, academic institutions, funding agencies, statistical agencies, media organizations, laboratories, healthcare, business, banking, industry
Environment (digital)	Search engines, recommendation systems, social networks, social media, e-businesses, data archives, institutional repositories, digital libraries, virtual organizations, cloud-based systems and services, mobile computing providers, information providers, data aggregators, data commons, data centres, virtual collaboratories and observatories

Organizational context	K-12, high education, public, for profit, non-profit, government, personal
People: Individual context	Home, workplace, virtual environments, Students, teachers, academics, researchers, information scientists and professionals, data scientists, computer scientists, software designers and developers, general public (shoppers, gamers, readers, searchers, users)
Embodiment/tangibility	Physical and tangible artifacts (robots, driverless cars etc.) Intangible (algorithms, agents, training data etc.)
Time and chronology	Historical, chronological, longitudinal developments and trends (1950s to date), Alan Turing test, first AI program, computer chess, expert systems, driverless cars, future AI

Table 5. A faceted taxonomy of AI literacy (table by author)

While it may be argued that the proposed high-level facets do not represent mutually exclusive categories, effort was made to ensure that the nuanced, subtle, and intangible aspects of AI literacy were captured. For instance, the *Embodiment/tangibility* facet is particularly important to include since numerous AI artifacts and applications on the web make use of codes, algorithms, and data that are invisible to a typical user. Similarly, the facet of *Ethical, legal, social aspects, issues, concerns* covers a range of key principles and theoretical constructs that may not be immediately apparent to or well understood by a user of an AI-enhanced technology. Furthermore, we argue that any taxonomy or knowledge structure that aims to map the emerging domain of AI literacy should take into account and build on the prior research, scholarship, and development in the areas of information literacy, digital literacy, data literacy, and internet literacy.

The proposed taxonomy fills the gap in existing AI taxonomies in a number of ways. It:

- Provides a unique AI literacy perspective that also draws upon previous research on information literacy, digital literacy, and Internet literacy
- Situates AI in the context of AI literacy education that emphasizes the importance of contextual as well as conceptual nuances of developing critical skills in learning about and use AI in its current context
- Stresses the interdisciplinary nature of AI and its implications for various educational and societal contexts, including everyday life and work
- Delineates the physical and digital nature of AI developments and their impact on, not only digital and intangible entities, but also on physical and tangible artifacts.
- Places particular emphasis on the notion of data as a key facet and component in all AI systems and technologies and its central role in ethical and responsible treatment humans in relation to AI systems.
- Draws upon and benefits from existing AI taxonomies, but enhances and expands their domain using published literature on AI literacy in various subject domains and disciplines, including computer science and more specifically artificial intelligence, information science, education, literacy research, and facet analysis and classification theory.

As Bawden (2001) argued, for the current digital information environment we need to develop a broad and complex type of literacy that would encompass all types of skill-based competencies without being confined or attached to one particular technology or set of technologies. This approach will ensure the technology literacy skills, knowledge, and competencies developed to date contribute organically and meaningfully to the development of a comprehensive AI literacy framework. It is worth reiterating Stordy

(2015) who argues that different types of literacies related to information, data, and digital technologies are complementary. The examination of research on AI literacy to date also confirms that different types of literacy are complementary rather than being mutually exclusive. It is, therefore, argued that AI literacy could be categorized under what Stordy (2015) refers to as ideological-peripheral perspective, i.e. the type of literacy that stems from new social practices as a result of the emergence of new technologies.

6. Conclusion

This paper proposed a taxonomy of AI literacy based on an examination of the existing literature on AI taxonomies, AI literacy publications, prior research on data/information/digital literacy research, and a metadata record analysis of AI literacy publications. It is challenging to create a comprehensive taxonomy for an emerging and conceptually complex domain such as AI literacy as it has to capture and represent two dynamically evolving and changing concepts, namely AI and literacy. However, given the urgent importance of learning and critical application of AI concepts, knowledge, competencies, and skills in this exponentially changing digital landscape, it is particularly timely to actively engage in the development of a conceptual framework for AI literacy. The taxonomy proposed in this paper is one of the first attempts at creating a concept map of the AI literacy domain in order to capture its different facets, subjects, and terms. It offers both theoretical and practical implications for the emerging area of AI literacy. Theoretically, it offers a conceptual framework for thinking and talking about what constitutes AI literacy. Practically, such a tool may serve to support AI literacy education and research in various educational, research, development, and organizational contexts. It provides a concrete framework upon which curricula for different target audiences can be developed. Given the high level and faceted nature of this proposed taxonomy, it has the potential to be further expanded and customized for various user communities and educational contexts. Furthermore, with the rapidly evolving generative AI and large language models and their widespread applications in many different disciplines, domains, and sectors, AI literacy taxonomies, knowledge maps, and subject categories have the potential to provide a conceptual foundation to support the critical understanding, use, application, and examination of AI-enhanced tools and technologies. This taxonomy offers a knowledge mapping and knowledge organization framework for educators, scholars, researchers, and the general public to imagine and develop the ways in which understanding and learning about AI in everyday life and in education can be conceptualized. The ultimate goal is to prepare and equip various user communities with knowledge, skills, and competencies that are required to critically understand, examine, and make use of AI tools and technologies Future research and emerging AI literacy frameworks and curricula will provide empirical, evidence-based, and user-centred data to support the creation of a more detailed taxonomic structure for the area of AI literacy.

References

Aitchison, J., Bawden, D., & Gilchrist, A. (2003). *Thesaurus construction and use: a practical manual*. Routledge.

Albrechtsen, H. (2015). This is not domain analysis. *KO Knowledge Organization., 42*(8), 557-561.

Bawden, D. (2001). Information and digital literacies: a review of concepts. *Journal of documentation, 57*(2), 218-259.

Broughton, V. (2006). *Essential thesaurus construction*. Facet Publishing.

Buchanan, B. G., Eckroth, J., & Smith, R. (2013). A Virtual Archive for the History of AI. *AI Magazine*, *34*(2), 86-86.

Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, *63*(2), 123-134.

Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, *10*(1), 29.

Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. *Business and Professional Communication Quarterly*, *86*(3), 257-295.

Erwin Jr, R. W. (2015). Data literacy: Real-world learning through problem-solving with data sets. *American Secondary Education*, 18-26.

Keene, J., Colvin, J., & Sissons, J. (2010). Mapping Student Information Literacy Activity Against Bloom's Taxonomy of Cognitive Skills. *Journal of information literacy*, *4*(1), 6-20.

Kong, S. C., Cheung, W. M. Y., & Zhang, G. (2021). Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence*, *2*, 100026.

Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in artificial intelligence*, *4*, 622364.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, *41*(4), 212-218.

La Barre, K. (2010). Facet analysis. *Annual review of information science and technology*, *44*(1), 243-284.

Lankshear, C., & Knobel, M. (2007). *Sampling" the new" in new literacies*. na.

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 100101.

Laupichler, M. C., Aster, A., & Raupach, T. (2023). Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy. *Computers and Education: Artificial Intelligence*, *4*, 100126.

Lee, J. H., Karlova, N., Clarke, R. I., Thornton, K., & Perti, A. (2014). Facet analysis of video game genres. *IConference 2014 Proceedings*.

Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16).

Martınez-Plumed, F., Loe, B. S., Flach, P., O hEigeartaigh, S., Vold, K., & Hernández-Orallo, J. (2018). The facets of artificial intelligence: A framework to track the evolution of AI. In *International Joint Conferences on Artificial Intelligence* (pp. 5180-5187).

McCarthy, J. (2007). What is artificial intelligence. Available from https://cse.unl.edu/~choueiry/S09-476-876/Documents/whatisai.pdf

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021a). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041.

Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., & Qiao, M. S. (2021b). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*, *58*(1), 504-509.

Niu, J., Tang, W., Xu, F., Zhou, X., & Song, Y. (2016). Global research on artificial intelligence from 1990–2014: Spatially-explicit bibliometric analysis. *ISPRS International Journal of Geo-Information*, *5*(5), 66.

Oeldorf-Hirsch, A., & Neubaum, G. (2021). What do we know about algorithmic literacy? The status quo and a research agenda for a growing field.

Ranganathan, S. R. (1967). *Prolegomena to library classification*. Asia Publishing House (New York).

Reddy, P., Chaudhary, K., & Hussein, S. (2023). A digital literacy model to narrow the digital literacy skills gap. *Heliyon*, *9*(4).

Ridley, M., & Pawlick-Potts, D. (2021). Algorithmic literacy and the role for libraries. *Information technology and libraries*, *40*(2).

Riel, J., Christian, S., & Hinson, B. (2012). Charting digital literacy: A framework for information technology and digital skills education in the community college. *Available at SSRN 2781161*.

Rizvi, S., Waite, J., & Sentance, S. (2023). Artificial Intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. *Computers and Education: Artificial Intelligence*, 100145.

Samoili, S., Cobo, M. L., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence.

Sandy, H. M., Froehlich, H., Hudson-Vitale, C., & Adkins, D. (2019). Topic Modeling and Facet Analysis of an Emerging Domain: Research Data Management and Data Curation. *NASKO*, 63-76.

Schuff, D. (2018). Data science for all: a university-wide course in data literacy. *Analytics and data science: Advances in research and pedagogy*, 281-297.

Schüller, K. (2022). Data and AI literacy for everyone. *Statistical Journal of the IAOS*, *38*(2), 477-490.

Shiri, A. (2014). Making sense of big data: A facet analysis approach. *KO KNOWLEDGE ORGANIZATION*, *41*(5), 357-368.

Stordy, P. (2015). Taxonomy of literacies. *Journal of documentation*, *71*(3), 456-476.

Stolpe, K., & Hallström, J. (2024). Artificial intelligence literacy for technology education. *Computers and Education Open*, *6*, 100159.

Street, B. V. (1984). *Literacy in theory and practice* (Vol. 9). Cambridge University Press.

Tenório, K., Olari, V., Chikobava, M., & Romeike, R. (2023, March). Artificial Intelligence Literacy Research Field: A Bibliometric Analysis from 1989 to 2021. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (pp. 1083-1089).

Vitolo, T. M., & Coulston, C. (2002). Taxonomy of information literacy competencies. *Journal of Information Technology Education. Research*, *1*, 43.

Wang, B., Rau, P. L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, *42*(9), 1324-1337.

Zaharee, M. (2013). Building controlled vocabularies for metadata harmonization. *Bulletin of the American Society for Information Science and Technology*, *39*(2), 39-42.

Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37.

Artificial intelligence (n.d.). In *Wikipedia*. Retrieved March 11, 2024, from https://en.wikipedia.org/wiki/Artificial_intelligence