

Your Data Body

Marilene Oliver
University of Alberta
Canada
marilene@ualberta.ca

Scott Smallwood
University of Alberta
Canada
Scott.smallwood@ualberta.ca

This paper delves into the aesthetics and ethics surrounding the collaborative virtual reality artwork, *Your Data Body*. Made using medical scan data as a metaphor for our ever-expanding bodies of intimate personal data, *Your Data Body* seeks to challenge how we interact with the data of others, questioning the etymology of the word data, meaning “given” and questioning whether in many cases, data is rather “taken”. Using the gaming device of moving through progressive scenes, users first encounter open-access anonymized scan data and later donated data given with active and ongoing consent of the subject. Each scene situates the medical scan data within LiDAR scans, is accompanied by poetic elements, and has a complex sonic composition that combines field recordings, choral composition and data sonification as a way to situate the data geographically, temporally and emotionally.

Virtual reality. Medical scan data. Personal data. Data privacy. Sound. Composition. LiDAR scanning. Data sonification.

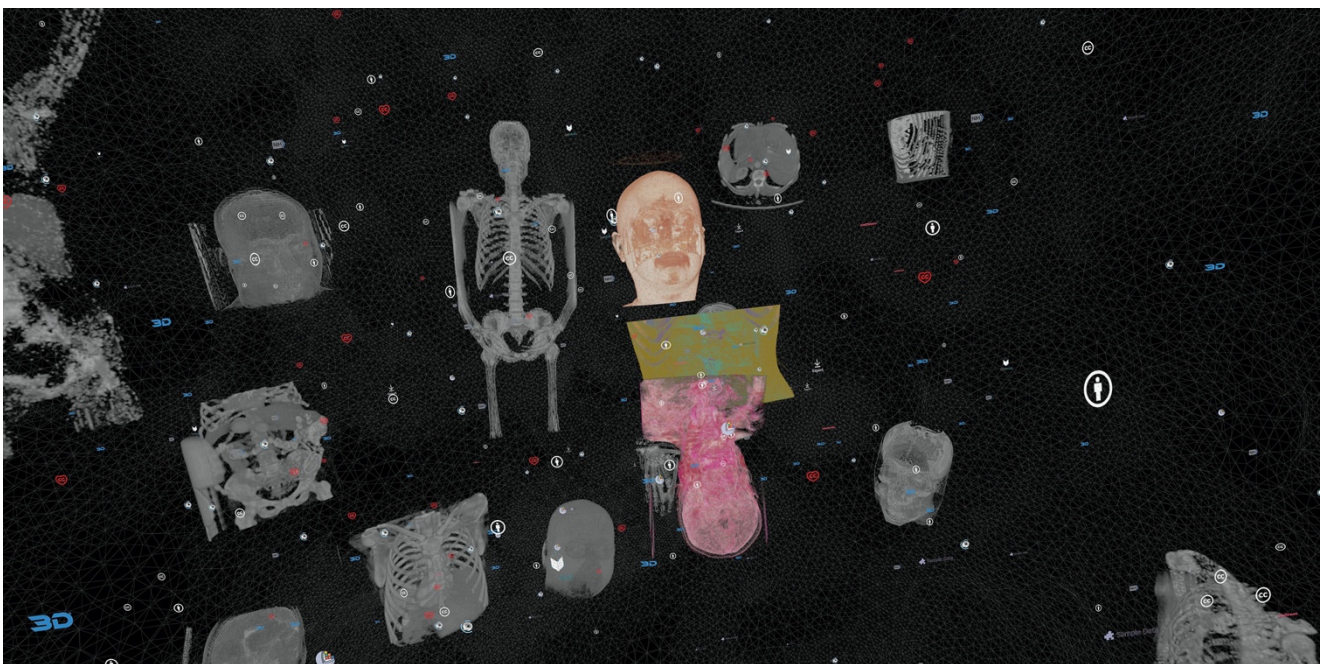


Figure 1: *Your Data Body*, scene 1.

1. INTRODUCTION

We are constantly warned that our personal data is vulnerable. We are told that it is used and abused by artificial intelligence, giant tech corporations, and controlling governments (Bridle 2018, O’Neil 2016). But do we really understand what “our data” consists of and what can be done with it by both ourselves

and others? Is it possible to unravel the complex entanglements of data gathering and automated processing technologies in order to see and understand what sociologist Deborah Lupton terms our “human data assemblages” in meaningful ways? Can virtual reality (VR) be used as a creative space to explore and situate our data bodies temporally, geographically and emotionally?

The research project Know Thyself as a Virtual Reality (KTVR) based at the University of Alberta, Edmonton, Canada has focused on these questions since 2019 through a series of interdisciplinary symposia, publications, and research-creation projects. The central research creation project is a diptych of VR artworks titled *My Data Body* and *Your Data Body*. *My Data Body* focuses on the data we generate and are responsible for as individuals, whereas *Your Data Body* questions how we interact with, understand, and are responsible for the data of others. Both the works use high resolution volume rendered medical scan data (such as MRI and CT scans) as well as other personal data, LiDAR scans, poetry and unique sonic compositions to create immersive and affective VR experiences that invite the user to think deeply about the data bodies they are seeing, hearing and touching, albeit mediated by a VR headset and controllers (Diodato 2012).

This paper focuses specifically on *Your Data Body*, the most recent of the two projects. As with *My Data Body*, medical scan data is used in the project both as a literal and metaphorical symbol of intimate personal data. Unlike *My Data Body* (Oliver et al. 2022) however which uses MRI scan data of the artist acquired especially for the work (Figure 1), *Your Data Body* is made using a combination of open-access anonymized datasets and donated medical scan datasets with varying levels of information about, and consent from, the subject of the scan. *Your Data Body* employs the gaming device of having progressive levels that the user moves through as a way to think through different levels of consent and data ownership, starting with open-access anonymized data and ending with data donated with active and ongoing consent and authorship. The different levels also situate the data within increasingly intimate and emotive virtual spaces, from the artist's home office to a series of old log cabins from a family property belonging to the scan subject. This paper will explain the aesthetic choices made in each of the scenes and share the many social, ethical and emotional questions that were raised by working with the personal data of others in VR.

2. SCENE 1: ANONYMIZED OPEN-ACCESS DATA

The first scene in *Your Data Body* collects open-access medical scan data that was originally acquired for scientific and medical research but has now been anonymized, so it can be used for secondary research with creative commons licensing. Since personal information has either been removed or obscured, consent from the scan subject is not required for use of these anonymized datasets. For *Your Data Body* computer tomography (CT), magnetic resonance imaging (MRI) and

positron electron tomography (PET) data were easily and freely downloaded from multiple research open access databases; The Cancer Imaging Archive (TCIA), National Library of Medicine, Embodi3D and Open Neuro. Additionally, openly available sample data for working in open-source radiology software (3D Slicer, OsiriX-Viewer and medDream) was also downloaded.

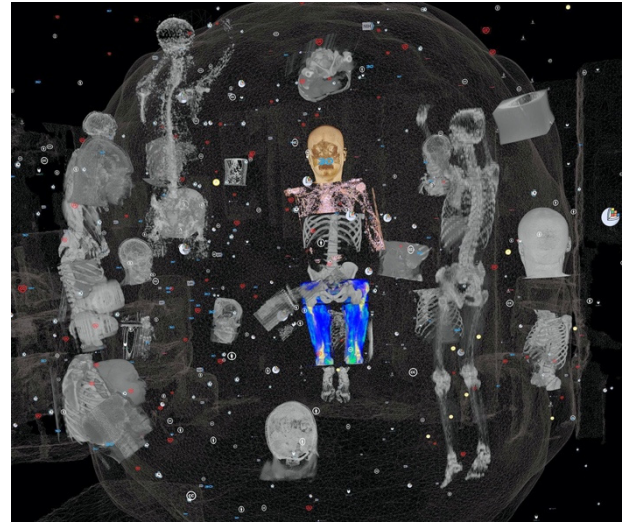


Figure 2: *Your Data Body*, scene 1.

Once downloaded, the scan data was processed so that it could be imported into VR, where it was placed in a web like pod, much like insects caught in a spider's web (Figure 2). The pod was then nested within a LiDAR scanned mesh of the home office in which the VR project was largely developed using the gaming software Unity. In early iterations of the scene, the data was placed in what looked like a storage facility or museum display case, but these structures felt strangely impersonal, resembling a science fiction computer game. In an attempt to acknowledge that it is the artist/researcher who is ultimately responsible for treatment of the anonymized data, as well as present where the data was being downloaded, transformed and rendered, it was decided to self-consciously place the data within a scan of the office in which the work was being developed. A medical data (DICOM) loader script was written for the gaming software Unity so that using the VR controllers grabbing function, the user can pull scanned body parts out from the web structure, place them elsewhere in the scene, resize and recolour them. The invitation is to use the scanned body parts like building blocks to create a Frankenstein like figure. As the user brings the scanned body part closer to them, and then to their ear like listening to a shell, they hear an automated voice reciting the information about the original research project and why the data was originally acquired. Each body part has a different automated voice, but the information about each of the datasets is very scientific and objective, giving no information

about the subject of the scan. Datasets which don't have this information are silent. An additional layer of sound in the scene is an abstract sonification of the data usage agreements for each of the databases (which range in length from a few lines to 18 pages long). Each body part also emanates logos and icons from them like confetti, which relate to the institutional or commercial database identity and the usage agreements attached to them.

The intention with this first scene is to invite users to reflect and question their relationship to open-access personal data. When the user is in the first scene of *Your Data Body*, they are able to hold the virtual head, chest, torso of another person in their own virtual hands. The body part is to scale so the user can relate to it with their own body, they can manipulate it and combine it with another anonymous body parts. But they know little about the subject, where or when the data was acquired, where the subject is now and of course the subject has no idea that a virtual copy of themselves is being held by another person some-where, somewhen in the world. Although arguably more uncanny and more embodied, this echoes many of our virtual interactions on social media where we consume, comment, like/love posts from strangers sharing very inti-mate information and images of themselves from every-where and everywhen. Or, as another example given by Laurence Scott in *The Four-Dimensional Human*, how platforms such as Airbnb allow intimate access into the interior home spaces of millions of strangers around the world at a time when we are increasingly socially distanced from our 'real' neighbours (Scott 2015).

The ethical difference of course is around consent and hopefully most people who share on social media and on Airbnb are doing it somewhat knowingly and actively. As explained earlier, consent is not legally required for anonymized data as the data has been de-identified and cannot be connected to the subject. Anonymization however has become less reliable in an age of big data, smart devices, and social media and anonymized datasets no longer offer the protections they once did. Smart devices and social media make a wealth of information publicly available [Cooper 2020, Parks 2021] and this big data can undermine the methods for protecting human subjects represented in anonymous datasets. Furthermore, as machine learning algorithms work by finding patterns in data, there is no assurance (or even way of knowing) if anonymized datasets are cross-referenced and thus reidentified. In order to test the theory of cross-referencing, a volume rendering of a team member's

own MRI scan was uploaded to Meta and it was tagged instantly (Figure 3). The potential for harm to the re-identified data subject is not equal for all data subjects and depends on the kind of data, its age, and a myriad of other local social, legal, political, and economic factors. The possibility of reidentification is particularly troubling when we consider that social media giants such as Google and Meta are establishing lines of business in the health care domain. For example, 23andMe is a Google venture (via Alphabet) that became a publicly traded company in early 2021 and is now using the DNA from millions of Americans to produce pharmaceuticals (Brown 2021). 23andMe is part of a Google health portfolio that includes insurance companies, medical record apps, and home health monitoring technologies that collect biometric data. Google's Project Nightingale, which gave Google access to health care data through research partnerships, has already raised privacy concerns and lawsuits (Schneble 2020). A solution developed in the radiological community for the re-identification of scan data is "defacing", which literally involves cutting away the face from head scans (Parker 2021). Defacing is mandatory for a dataset to be uploaded to the Open Neuro database, but as of writing, not to Embodi3D, nor 3D Slicer, other open access scan databases.

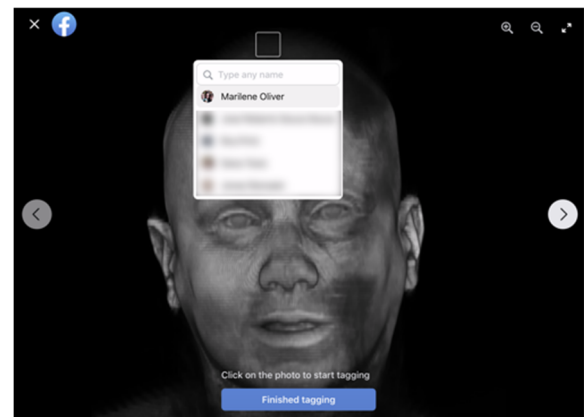


Figure 3: *Your Data Body*, scene 1.

Again, there is considerable debate over the efficacy and ethics of defacing. Many however are horrified when they first encounter a 3D rendering of a defaced dataset for it resembles a head with its face violently axed off. Aesthetically, a defaced 3D volume rendered dataset brutally symbolizes the crude (and typically automated) attempts to de-personalize and de-humanization data (hence the choice to include it in this scene (Figure 4)).



Figure 4: *Your Data Body*, detail of defaced data

In an attempt to re-humanize and re-personalize the data in this scene and indeed the whole of the *Your Data Body* artwork, the project composers wrote a choral composition that the user is enveloped by when they are within the data nest. The composers worked with a university choir to make a multichannel recording of the choir chanting “Anon, Anon, Anon.” This layer of beautiful harmonious human voices is intended as a gratitude to all the anonymous data that makes so much contemporary research and medical advancement possible. As the sound is so crucial in *Your Data Body*, the sound engine WWise, was integrated into Unity to allow more spatial control of the data sonification, automated and human voices.

3. SCENE 2: RINGS OF FAMILIARITY AND COLLECTIVE DATA OWNERSHIP



Figure 5. *Your Data Body*, Scene 2

The next level or scene in *Your Data Body* is made up of MR and CT scans “donated” or “given” to the project specifically for creative purposes. Following academic research ethics procedures, an open call for scan data was disseminated through various academic email lists, as well as through direct requests to friends and family. In “Consent to Our Data Bodies: Lessons from Feminist Theories to Enforce Data Protection,” Paz Peña and Joana Varon bring a feminist lens to thinking through consent to data usage (Peña and Varon 2019). They explain the problematics of binary consent options

and the illusion that consent can be a free, rational, and individual choice. When a simple click can give access to a website, or, in the case of some medical research projects, to a potentially life-saving study, is there really a choice? In the same way that terms and conditions for social media sites are unreadable, is there a risk that research ethics information sheets and consent forms are equally unintelligible, prompting us to click through rather than meticulously scroll through the entire text? Peña and Varon suggest that the act of consent needs to be a) active, meaning actively agreeing with body and words to do so (not only the absence of no); b) clear and intelligible; c) informed, fully conscious; d) freely given, out of choice and free will; e) specific to a situation, therefore f) retractable and g) ongoing. With these healthier qualifiers of consent in mind, those who donated their data were shown the project during its many stages and given the option of retracting their permission or suggesting changes and being credited as an author.



Figure 6. Screenshot of Scene 2 showing text falling

The intention in the second scene is to suggest a collective and consensual ritual of offering and receiving data. When the user enters the scene, they are surrounded by rings of data that swirl around them and they are beneath a virtual sculpture of an eight-armed figure hanging from above seemingly reaching to catch or release data (Figure 5). There are three concentric rings of swirling data around the viewer. Data from close family and friends is in the first ring, then known colleagues in the second ring, and less familiar colleagues of colleagues or friends of friends in the outer ring. The rings float within a 360-degree video sphere of a forest of trees shedding their autumn leaves captured close to where the VR artwork was made, signifying being in a time of harvest. Inside the dome are various LiDAR scans of circular structures; a topiary trained tree, a stone circle and a large lily pond. These structures were all scanned by the lead artist in locations near their father's home and chosen because of their relationship to time, tradition and ritual. From above fall words from another chant like poem called Touch written by J.R. Carpenter about the desire for embodied connection in a virtual world (Figure 6). If the user grabs/touches any of the data, it starts to

glow with a warm pulsing light. As the user moves closer to the scans, they hear human voices either reciting or singing the Touch poem that then wash through them as the scans swirl around the scene. The people who donated the data were given the option of either reading/singing the poem themselves in their own voice, or allowing a trained choral voice or musical instrument to be attached to their data. Most of the people who donated the data did agree to record in their own voice, but made the recordings in different places and at different times. In VR however the user is surrounded by a circular chorus of the scan subjects reciting or singing the poem at the same time. The idea of placing the data in concentric circles around the viewer developed from learning about examples of collective, rather than individual concepts of data ownership. In “Neuroethics Questions to Guide Ethical Research in the International Brain Initiatives,” neurotech ethicist Karen Rommelfanger and her colleagues explain how work done by International Brain Initiative, to coordinate brain mapping initiatives in China, Japan, Korea, Australia, and the EU, highlighted important differences between Western and Buddhist and Confucian societies in relation to data collection and ownership (Rommelfanger 2018). Rommelfanger explains that as China, Japan, and Korea societies are typically more collectivist, medical decisions and ownership of a person’s data is collective and family-based rather than individualistic.

Katherine Hayles famously argued that the tech industry continually revels in the fantasy of a disembodied future where we can cast our feeble bodies and become ostensibly omnipotent and omnipresent bytes of information (Hayles 1999). But our data is hardly objective and acts instead like an expression of what Donna Haraway called our situated knowledges. It is full of the same biases, privileges, limitations, and historical locatedness as our physical bodies; no matter how much one might want to dissolve into nothingness, there are always traces that lead us back to our bodies in the here and now (Haraway 1988). This ‘limitation’ is not only inescapable, but the source of our hopes and fears, individuality, and sense of collectivity. VR offers a unique opportunity to challenge this post humanist logic by showing how our personal data continues to be sensorial, affective and highly relational, offering the ability to be sur-rounded by and immersed in it.

4. SCENES 3, 4 & 5: SITUATING DATA AESTHETICALLY, PHYSICALLY AND EMOTIONALLY

One participant responded to the call for scan data with a donation of over twenty datasets is herself an artist. The artist has created several artworks with her own medical scans, which have been acquired

for diagnostic reasons since 2014 as part of her ongoing oncological care. In 2019 for instance, she worked with her husband and long-term collaborator to create *Light Touch*, a large silk fabric tent printed with images of her brain scan held tenderly in both her own and her husband’s hands. In this and subsequent works, the artist and her husband have made strong aesthetic choices about how the artist’s scans are presented; typically fragile, transparent, and intermingled with images of flowing water and poetic text. Through conversations and experimentation, it became clear that the artist who donated the data and her husband should collaborate on the *Your Data Body* project, be offered aesthetic control of the scenes that feature her data (3, 4 & 5) and become collaborators in the work.



Figure 7. Screenshot of Scene 3, *Your Data Body*



Figure 8. Screenshot of Scene 4, *Your Data Body*

The final three scenes of *Your Data Body* situate the artist’s medical scans within LiDAR scans of disintegrating pioneer wood cabins in the boreal forest of Alberta, a location that has been central to the artist’s life past and present as a beloved second home. Originally a homestead property and then a fishing camp, the location is now a collection of wooden cabins in varying states of decay and repair as the artist and her family are committed to ‘rewild’ the proper-ty. In the third scene, rows of the artist’s

head scans sit within the ruins of the “homestead” cabin (Figure 7). In the fourth scene, a CT scan of the artist’s chest and lungs rotate slowly within the “Miller” cabin (Figure 8), and in the fifth and final scene, the artist’s fiery PET scan data is cradled within the “tent” cabin (Figure 9). In each scene is a poem written by the artist’s husband that either floats above her data, is nestled within it, or as in the final scene, sways through it back and forth. Her husband wrote after first seeing renderings of his wife’s data and feeling a deep sense of lack in them, a lack of recognition, a lack of ownership, a lack of meaning. The poems are tender, mournful, sensuous and full of longing, recalling shared intimate memories of each other’s bodies in a very specific location steeped in personal and socio-political history.

Once again, sound is crucial in these scenes for which the composer went to the lake site and made a library of field recordings of water, fire, wind in trees, birds and animals, and distant trains and traffic for the project. The composer also used the cabins and found objects as instruments, capturing their reverberances. With their students and working with WWise, the composer placed the captured sounds in the scenes to guide and envelop the user. Inside the rotating chest of the Miller cabin is the sound of water (Figure 8), in the cradled PET scan, the sound of fire (Figure 9). In other parts of the scene are recordings of the forest and the lake. In addition to the spatial composition of field recordings, there is also a layer of choral singing and spoken word based on the husband’s poems. To the right of the cabin which holds the PET scan for instance, there are scanned trees in which whisperings of poems can be heard.

When reflecting on their experience of making these scenes, both the artist and her husband said they found the process surprisingly therapeutic and positive. What once had been scans that conjured fear and anxiety, now felt loved and cared for. Working with the scans in VR and making aesthetic decisions about them gave them agency and ownership of their data. Interesting, having worked for so long in virtual reality with the medical data within the LiDAR scans of the cabins, when they now return to the cabins in real time and space, they almost feel the data within it.

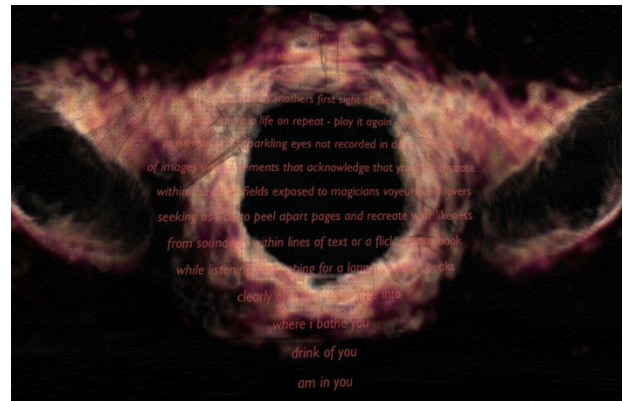


Figure 9. Screenshot of Scene 5, *Your Data Body*

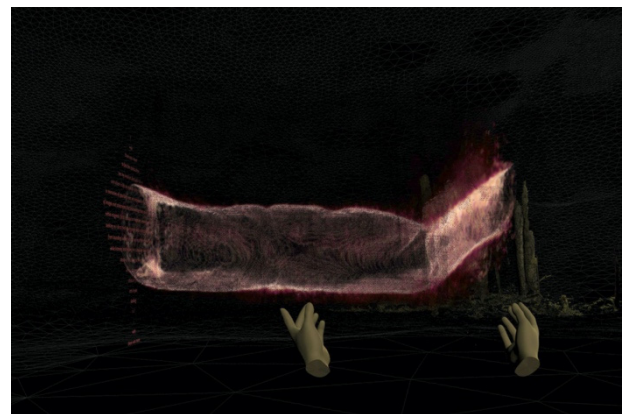


Figure 10. Screenshot of Scene 5, *Your Data Body*

5. REFERENCES

- Bridle, J. (2018) *New Dark Age: Technology and the End of the Future*. Verso, UK.
- Brown, K. (2021). 23andMe to Use DNA Tests to Make Cancer Drugs: All Those 23andMe Spit Tests Were Part of a Bigger Plan.” Bloomberg Business, November 4, 2021.
<https://www.bloomberg.com/news/features/2021-11-04/23andme-to-use-dna-tests-to-make-cancer-drugs> (retrieved 20 December 2023).
- Cooper, A., and Coetzee, S. (2020) On the Ethics of Using Publicly-Available Data. In eds. Hattingh, M, Mat-thee, M, Smuts, H, Pappas, I, Dwivedi, Y. and Mäntymäki, M. *Responsible Design, Implementation and Use of Information and Communication Technology*. Springer International Publishing, New York.
- Diodato, R. (2012) *Aesthetics of the Virtual*, State University of New York Press, New York.
- Haraway, D. (1988) Situated Knowledges: The Science Question in *Feminism and the Privilege of Partial Perspective*. In *Feminist Studies* 14, no. 3, 575–99.

- Hayles, K. (1999) *Toward embodied Virtuality. In How we became Posthuman: Virtual bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press. Chicago, US.
- Lupton, D. (2020) *Data Selves: More-than-Human Perspectives*. Polity Press, Cambridge, UK.
- Oliver, M. Scott, S. Carpenter, J.R., Moore, S. (2022) My Data Body. In: Bowen, J.P., Weinel, J., Borda, A., Diprose, G. (eds) (2022) *EVA London 2022: Electronic Visualisation and the Arts*. doi:[10.14236/ewic/EVA2022.7](https://doi.org/10.14236/ewic/EVA2022.7)
- O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data increases Inequality and Threatens Democracy*. Crown, New York, US.
- Parks, A. (2021) Unfair Collection: Reclaiming Control of Publicly Available Personal Information from Internet Data Scrapers." SSRN Scholarly Paper ID 3866297.
<https://papers.ssrn.com/abstract=3866297> (retrieved 20 Dec 2023).
- Parker, W, Jaremko, J, Cicero, M Azar M, El-Emam, K. Gray, B, et al. (2021) Canadian Association of Radiologists White Paper on De-Identification of Medical Imaging: Part 2, Practical Con-siderations in *Canadian Association of Radiologists Journal*, 72 (1), 25–34.
- Peña, P, and Varon, J. (2019) Consent to Our Data Bodies: Lessons from Feminist Theories to Enforce Data Protection.
<https://codingrights.org/docs/ConsentToOurDataBodies.pdf> (retrieved 20 Dec 2023).
- Rommelfanger, K. Jeong, S. Ema, A. Tamami Fukushi,T. Kasai, K. Ramos,K, Salles, A and Singh, I (2018) Neuroethics Questions to Guide Ethical Research in the International Brain Initiatives in *Neuron*, 100 (1), 19–36.
- Scott, L (2015) *The Four-Dimensional Human: Ways of Being in the Digital World*. Penguin, UK.
- Schneble, O, Elger, B. and Shaw, D. (2020) Google's Project Nightingale Highlights the Necessity of Data Science Ethics Review in *EMBO Molecular Medicine* 12 (3): e12053.