Ethics for Digital Identity and Identity-Driven Algorithms

By Mike Kiser (Sailpoint)

To comment on this article, please visit our GitHub repository and submit an issue.

Table of Contents

Abstract

Recent work by multiple organizations has helped to solidify a practical approach to ensuring the ethical use of digital identity technologies, including identity-driven algorithms. This article seeks to elucidate one approach that can improve internal ethical posture and foment comparisons between existing implementations across various industries. It establishes five ethical axes to evaluate and communicate relative strengths and weaknesses between implementations. The five key axes asserted herein include (1) well-being, (2) accountability, (3) transparency, (4) fairness, and (5) user data rights. Finally, the article explores a methodology for applying those axes within an organization.

Introduction

Algorithms that rely on digital identity and other data from the past to predict the future have long been commonplace: they predict what we will watch next,[i] how financially stable we will be,[ii] or how likely we are to commit a crime.[iii] The assumption is that with enough data, anything is predictable. Over the last several years, however, headlines have repeatedly illustrated the influence of algorithms on human well-being and highlighted many inherent algorithmic biases.[iv] Before we embrace a digital identity solution or an associated algorithm, how can we ensure that they promote justice and fairness rather than reinforcing existing inequalities? How do we prevent new and unforeseen harms?

To answer the questions above, it is not enough to merely espouse ethical principles: "Do no harm" is only helpful if it comes with the tools for everyday use. The industry needs a practical and measurable standard for digital identity and identity-driven algorithms: this may enable improved ethical posture and the comparison between existing implementations of algorithms across various industries. This article looks to recent work by IBM,[v] the IEEE,[vi] and the High-Level Expert Group on Artificial Intelligence (AI HLEG)[vii] , which have helped to solidify an ethical approach to the use of identity-driven algorithms. While these efforts focus on algorithms, their principles can provide the foundation for an ethical approach to digital identity more generally.

This paper proposes five ethical axes that build off of the work cited above and enable comparisons between digital identity solutions:

1. Well-being
2. Accountability
3. Transparency
4. Fairness
5. User data rights.s

All three of the initiatives mentioned above share these five axes as core to an ethical approach to artificial intelligence are shared by, with varying emphasis on each of them. For example, the IEEE calls out the impact of effectiveness misuse and competence of a solution and provides a list of resources for further investigation. The AI HLEG provides a tiered model that describes how ethical principles translate into more practical activities. IBM's design emphasizes the importance of establishing ethics as an underpinning part of a solution's design rather than an afterthought.

The five axes chosen in this paper reflect common elements of these approaches; together, they form a proposed framework and basis for further ethical exploration by identity practitioners.

Note that IDPro is concurrently releasing another article, *Ethics and Digital Identity*, which adopts some of these values and adapts others. That article explores the theoretical underpinnings of ethics – including the importance of having conversations and developing a shared understanding of ethical values within the communities served. As such, we strongly urge you to consider the values herein as inputs to a conversation rather than the end of one.[viii]

A Proposed Ethical Maturity Graph

This article proposes a standard approach bolstered by easily accessible tools that create a practical ethical evaluation and comparison method. Any analysis or evaluation is only useful if its intended audience understands it. To that end, a radar chart is recommended to communicate the evaluation of any particular use of digital identity, artificial intelligence, and associated algorithms. This visual representation creates a means for evaluating past ethical progress and an implementation's relative strengths or weaknesses.

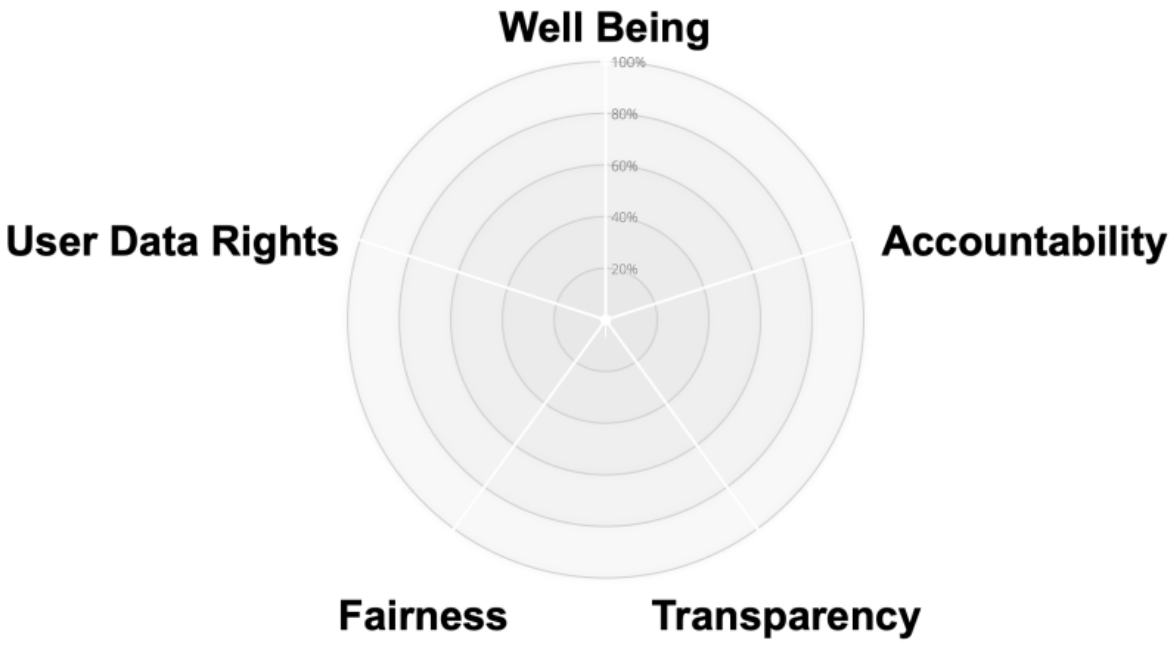

Figure 1 - A Proposed Framework for Ethical Evaluation

Each of the five axes in Figure 1 is assessed and then plotted on a radar chart for quick visual comprehension (the center of the graph connotes complete ethical immaturity for that axis).

Since every organization seeking to utilize technology ethically has unique requirements or needs, creating this graph involves self-evaluation. At the same time, it is important to note that practitioners may find common processes and tools helpful for evaluating each axis.

The following section examines each axis in Figure 1, highlighting key processes and best practices that will increase ethical maturity. Where possible, it offers open-source tools or readily available techniques and provides a concise guide for self-evaluation along that particular axis.

Well-Being

The first tenet of a practical ethical approach is that an identity practitioner must put human well-being first. This is not normally a controversial assertion. The question, then, is what the term "well-being" signifies. Certainly, there is an established, widely accepted truth that grounds well-being and is expressed in the Hippocratic oath: "Do no harm," but the concept must involve more than a mere defensive approach.

As explored in *Ethics and Digital Identity,* the question remains: whose well-being should be prioritized?[ix] Practitioners often make difficult choices: to protect one group or class may negatively impact another. Determining well-being is not simple when wielding the power of identity since well-being is nuanced and often culturally dependent. This context-dependence means exploring and documenting the interplay of interests for all parties. There may not be clear-cut answers for complex issues, but a methodological approach is essential for understanding the impact of decisions made in the development of new solutions.

Using a tool known as an "ethics canvas," these choices and outcomes may be explored and centrally documented.[x] Identifying the affected individuals, their relationships, and worldviews helps to give concrete insight into whose interests might be in conflict and what trade-offs underlie each decision or implementation choice.

Figure 2: Example of an Ethics Canvas

This process helps all participants — from designers to implementers — to contemplate the ramifications of their decisions and to embed a mindset that seeks to underscore the morality of their actions more clearly. This central document establishes the ethical standard that all participants agree to abide by; it is a guiding force throughout the rest of the process.

Figure 3: Human Impact in the Ethical Maturity Framework

This documentation increases ethical maturity by ensuring that the organization has, at minimum, explored and documented its human impact and, at best, regularly reviews whether its impact matches its intent.

Accountability

Once the ethical choices have been laid out using the ethics canvas above, organizations seeking ethical maturity must ensure that they are accountable for meeting the newly documented ethical standard. There are two primary ways to achieve accountability. First, all ethical decisions must be tracked as solutions that are designed and architected. Documentation encourages designers, architects, and implementers to make ethical choices and records a rationale for each choice. Second, a feedback loop must be built into

the solution so that the end users (those directly impacted by the technology) have a method for holding the designers and creators of identity systems accountable.

Documenting past choices reinforces the notion of responsibility for those decisions – something people are far too willing to abdicate. Previous studies have found that as technology emerges, humans become more reliant on the technology rather than their own faculties. An easy illustration of this occurred as mobile devices became popular: before the rise of the mobile phone, most people had at least a score of numbers committed to memory so that they could call friends and family. Subsequently, they relied on their devices — technology — to provide that number recall function. Research has shown that this process is writ large across society: technology such as search engines and the internet have shifted recall from the actual content (phone numbers, key facts, dates, and navigation) to the location to find that content online.[xi] In short, we have delegated that memory function to technology. There is an inherent danger that using identity data (particularly in algorithms) will lead people, organizations, and ultimately large swathes of humanity to cede ethical choices. They will merely remember "whom to ask" (or what service to call) rather than feeling accountable for the decisions themselves. Documenting choices compels technologists to remember that the choice (and its consequences) lies with them.

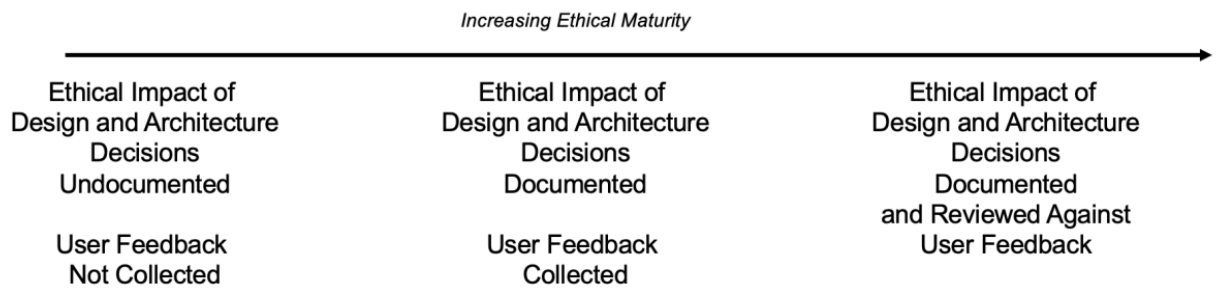

Figure 4: The Feedback Loop of Ethically Mature Organizations

Figure 4 shows that accountability is more than internal processing. Giving voice to those impacted by the use of identity data provides essential feedback, illuminating the actual effects of an implementation. Features that allow comments and complaints into our systems in ways that are easy to use are not optional for an ethically mature organization. Far from being an "add-on," this is a requirement for accountability — if the goal is fairness, then those who feel that a particular use of identity data or that an identity-based decision is unfair must be able to register their complaint.

Transparency

For users to hold practitioners accountable, the reasons for those decisions must be transparent. True transparency not only answers the question of why, but it breaks down the how in a simple way. When applying this to identity, it is critical to use language that is easily understandable to the end user, as technical jargon can quickly become complex.

Figure 5: Ethical Maturity in Relation to Transparency

This transparency is more than just disclosure for disclosure's sake — it builds trust in the power of identity itself. Mature identity systems that can explain themselves and invite feedback (for accountability) promote a "virtuous loop" that increases the likelihood of an ethical standard undergirding a particular identity implementation.

Fairness (Bias)

For identity to promote fairness, it must expose its own biases. Detection is successful only when the full range of bias is understood: there are helpful tools out there, such as IBM's taxonomy of bias,[xii] which ranges from "shortcut bias" (doing too much, too quickly) to "impartiality bias" (false assumptions of sound judgment) and more direct prejudices (like "self-interest bias").

Identifying bias is sometimes straightforward, such as when a flawed identity-based algorithm is blatantly biased, even to outsiders with little knowledge of the system. It is easier to spot and address these issues rapidly.

In other instances, biased outcomes can be more insidious: they must be discovered through diligent examination of the identity data and the process by which a system assimilates and learns from that data set. Hiring or salary determination algorithms that use historical data, which tend to depress the value of women in the marketplace,[xiii] provide a well-known example of this. These biased data sets reflect past sociocultural trends in which women earned less money due to institutional, structural, familial, and other drivers. This poor data becomes a self-fulfilling prophecy that locks women into the patterns of the past.

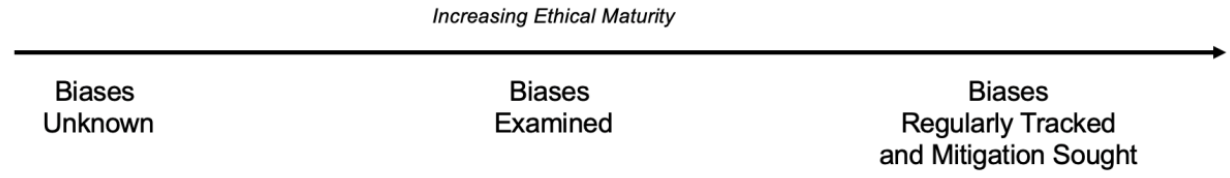

Figure 6: Ethical Maturity in Relation to Fairness

There are many examples in which algorithms can perpetuate biased outcomes. To achieve a mature ethical posture, organizations that identify one of these issues at play in their model or proposed implementation should return to the well-being axis: with the new information in hand, a reevaluation of the use of technology (e.g., AI or other algorithms) is necessary to understand the impact on affected parties–and to discern whether or not the system can be refactored to address and correct for inequality.

User Data Rights

User data rights underpin other essential ethical considerations. Rather than a nice-to-have, privacy and control over the data that comprise identity is recognized as a fundamental human right. Article 12 of the Universal Declaration of Human Rights, adopted in 1948 by the General Assembly of the United Nations, states, "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks."[xiv] While we may not have had a complete reckoning over what this means in the digital age,[xv] more than 160 countries have either enacted data privacy laws or are in the process of doing so – reinforcing that the right to control personal attributes and data is moving towards near-universal acceptance.[xvi]

Assuming that ethical maturity is the goal, technology that supports user data rights should be an automatic inclusion for anyone seeking to use identity and personally identifiable data. These include standards that grant the user control over the use of their data, such as enabled by User Managed Access (UMA), consent management, and selective disclosure. Furthermore, organizations should apply techniques that ensure personal data cannot be associated with individuals and correlated (data anonymization, pseudonymization, and differential privacy).[xvii] These techniques and tactics provide a firm grounding for realizing ethical standards.

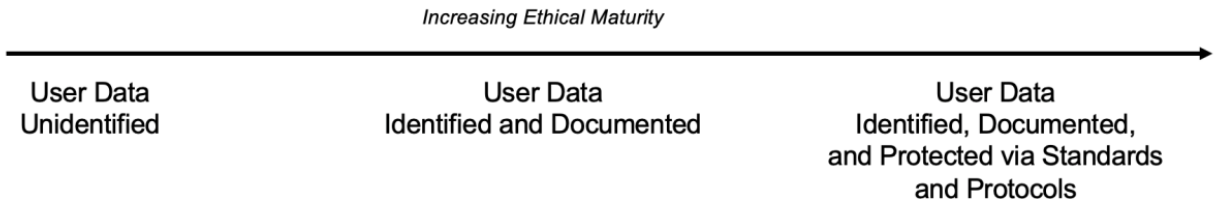

Figure 7: Ethical Maturity in Relation to User Data Rights

Ethically mature organizations understand what data they are collecting, justify how and why it is being used with respect to their ethical position, and take sufficient steps to protect that data.

Note that "user data" can go beyond individuals' private data. Even when users willingly consent to use their data – and even when data is collected anonymously – that data can

be used to do unethical things. Elizabeth Renieris notes in her book *Beyond Data* that an overreliance on "user control" can open the door to other harms without suitable controls.[xviii] These harms might include data uses that the user misunderstood or did not foresee (after they clicked "consent" on a screen that lacked transparency). It may also include attempts to use mass amounts of anonymized data to manipulate (or, in UN parlance, "arbitrarily interfere with") behavior. The Cambridge Analytica scandal provides a valuable example.[xix]

Evaluating Identity System Ethics

Once organizations carry out a self-audit on the axes above (or those they have selected after reflection and conversation), then the radial chart enables stakeholders to see the identity system ethics at a glance:

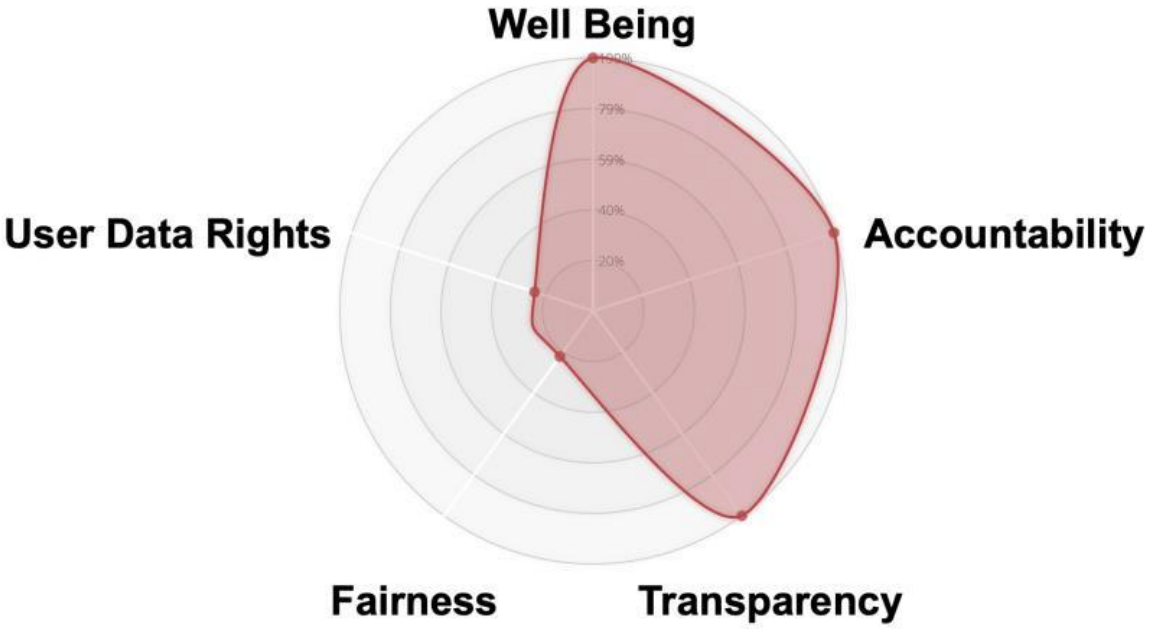

Figure 8: Example Ethics Analysis Radial Diagram

Measures can then be taken to address areas of weakness. In the figure above, well-being is a well-founded goal of the organization. Still, there are issues with fairness and a near disregard for the principles of protecting user data rights, which are likely revealed by the transparent system that enables accountability (since both of these score highly).

The goal is continuous improvement, such that we are:

- More aware of the impact of these systems on human well-being
- Working for accountability internally and externally
- Increasingly transparent in our decision-making

- Seeking out other voices that might go unheard
- Using standards that improve how user data rights are protected.

True ethical advancement requires, however, that the journey is not a solitary one. By comparing various algorithms, their data inputs and outputs, and how they are implemented, practitioners can begin to learn from one another. Lessons learned by one organization can spread rapidly throughout the community, advancing the cause of ethics and embedding an ethically-minded approach in the next wave of solutions.

Projecting several hypothetical analyses of comparable solutions on a single ethical maturity graph rapidly shows the relative strengths and weaknesses of each in an intuitive, visual way:

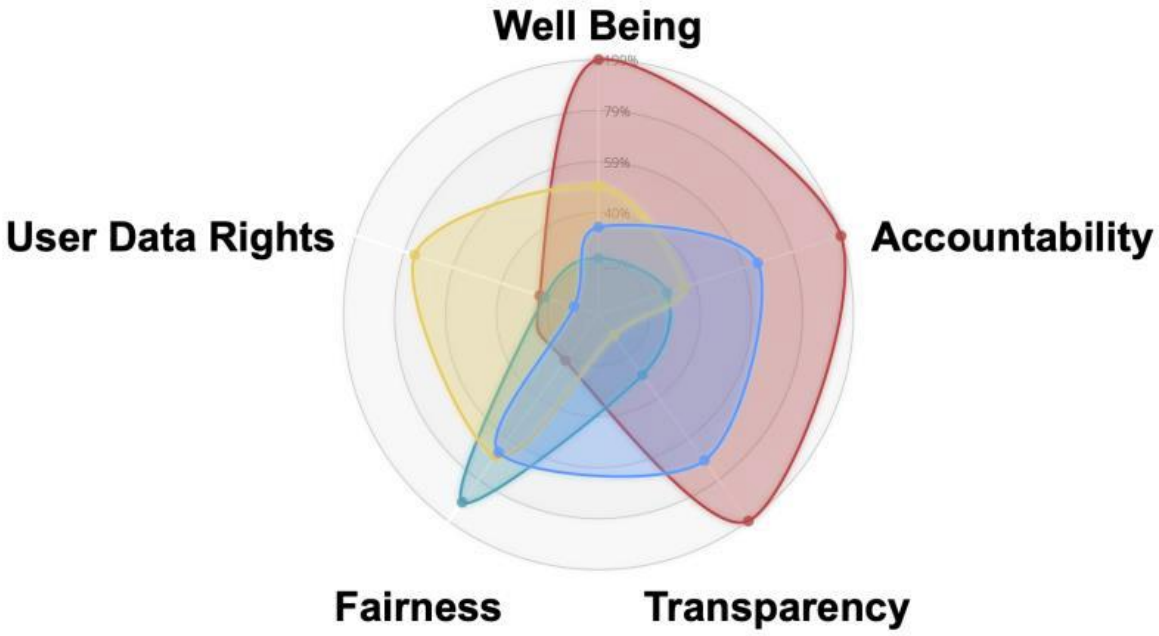

Figure 9: Comparing Implementations on Ethical Axes

Rather than abstract discussions that are often vague, a visual representation of the ethical status of an algorithm provides a framework to facilitate interaction and allow for productive learning between organizations and across an industry.

Conclusion

Establishing and following an ethical standard for using digital identity — before developing the technology— is essential. As we race to embrace new technology, mine new data sources, and that anything is predictable with enough data, we would do well to pause. As

identity systems and artificial intelligence converge, it is worth remembering what a Washington State Supreme Court case recently found:

"We affirm this court's long history of recognizing that one's past does not dictate one's future."[xx]

Past patterns in aggregate can be helpful, but we must not lose sight of the individuals involved in the systems we build: to ensure that people are treated ethically creates ripples through society at large.

This article has sought to further the practice of ethics within the digital identity industry by delineating a pragmatic and measurable ethical approach for those involved in designing, directing, creating, and administering the next wave of technology. This five-pronged approach to ethics promotes discussions between those seeking to take an ethical approach to all phases of development and thus advances human well-being through innovation.

Done well, we build trust in the small, everyday tasks that, in turn, promote equitable outcomes and build systemic trust. Lest we be seduced by the power and potential of technology, we must not allow it to outpace our ethics. In short, we must strive to use digital identity with our humanity intact.

---

[i] D. Jackson, "The Netflix Prize: How a $1 Million Contest Changed Binge-Watching Forver," 7 7 2017. [Online]. Available: https://www.thrillist.com/entertainment/nation/the-netflix-prize.

[ii] K. Waddell, "The Atlantic," 2 December 2016. [Online]. Available: https://www.theatlantic.com/technology/archive/2016/12/how-algorithms-can-bring-down-minorities-credit-scores/509333/. [Accessed 12 July 2019].

[iii] J. Angwin, J. Larson, S. Mattu and L. Kirchner, "Machine Bias," 23 May 2016. [Online]. Available: [1] ProPublica, "Machine Bias," ProPublica, 23-May-2016. [Online]. Available: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. [Accessed: 01-Aug-2019]. . [Accessed 12 7 2019].

[iv] See 2 and 3 above.

[v] A. Cutler, M. Pribić and L. Humphrey, "Everyday Ethics for Artifical Intelligence," IBM Corp, New York City, 2019.

[vi] "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems," 2019. [Online]. Available: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf. [Accessed 15 July 2019].

[vii] AI HLEG, "Ethics Guidelines for Trustworthy AI," European Commission, Brussels, 2019.

[viii] Marsman, H. (2024) "Ethics and Digital Identity" IDPro Body of Knowledge 1(14) doi: https://doi.org/10.55621/idpro.104.

[ix] Ibid. Note, in particular, the discussion of Utilitarianism and the value of well-being.

[x] "The Ethics Canvas," 2017. [Online]. Available: https://ethicscanvas.org. [Accessed 13 July 2019].

[xi] B. Sparrow, J. Liu and D. M. Wegner, "Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips," Science, vol. 333, no. 6043, pp. 776-778, 2011.

[xii] A. Cutler, M. Pribić and L. Humphrey, “Everyday Ethics for Artificial Intelligence,” IBM Corp, New York City, 2019.
[xiii] C. O'Neil, Weapons of Math Destruction, New York: Crown Publishers, 2016.
[xiv] United Nations, “Universal Declaration of Human Rights,”10 December 1948, https://www.un.org/en/universal-declaration-human-rights/.
[xv] See, for example Renieris, Elizabeth M. *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse*. MIT Press, 2023.
[xvi] D. Banisar, “Banisar, David, National Comprehensive Data Protection/Privacy Laws and Bills 2018 (September 4, 2018). Available at SSRN: https://ssrn.com/abstract=1951416 or http://dx.doi.org/10.2139/ssrn.1951416,” 4 September 2018. [Online]. Available: https://ssrn.com/abstract=1951416. [Accessed 12 July 2019].
[xvii] Kantara Initiative, “WG - User Managed Access,” site accessed 14 January 2020, https://kantarainitiative.org/confluence/display/uma/Home
[xviii] Renieris, E. (2023) “Beyond Data”.
[xix] See, for example, Confessore, N. "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far" New York Times: London (April 4, 2018)
[xx] https://www.courts.wa.gov/opinions/pdf/2016715.pdf [Accessed 22 July 2024].