Module 3 •

Data agency and sovereignty

Authored by:
Javiera Atenas •

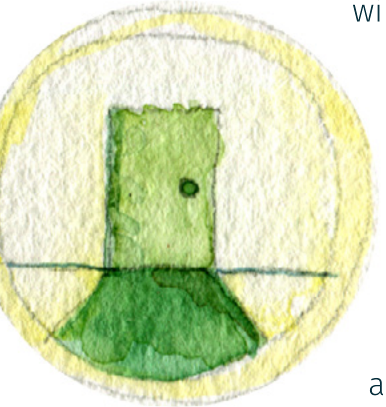
This project, Understanding data: praxis and politics is funded by the EPSRC and the Human Data Interaction Network+, under the grant number EP/R045178/1. **Caroline Kuhn H.** is the Principal Investigator and responsible for the project. •

How to cite this chapter: Atenas, J. (2021). Data agency and sovereignty. In C. Kuhn, J. Atenas & L. Havemann (Eds.), Understanding Data: Praxis and Politics. HDI - Data, Praxis and Politics. • https://doi.org/10.5281/zenodo.5137402. •

Module 3 • **Data agency and sovereignty •**

Content:

• Abstract and description of the module •



This module is aimed at presenting two-core elements of the world of data, one at an individual level and the other, at the collective level. In the first unit, we will review the concepts and skills needed to enable personal data agency, which can be understood as the individual's ability to understand and challenge the data collected about him/her. By doing so, he/she will be able to make informed decisions about his/her data through understanding the legal landscape of data protection and data rights. This will allow individuals to curate and control their (personal) data. To do this we need to gain knowledge of these data-driven systems, being able to identify them as well as understanding how they operate. In the second unit, we will present the key elements and principles of indigenous data sovereignty (ID-SOV). This is a relatively recent concept that can be understood as the right of indigenous peoples to own, control, access and possess data that derive from their needs and social reality. This is grounded on the rights to self-determination and governance as affirmed in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP).

• Learning outcomes •

1. Understanding the core elements of personal data agency

2. Understanding how to enable personal rights through personal data agency

3. Understanding the concepts of data agency and data sovereignty

4. Acquiring the abilities to manage and challenge personal and collective sensitive data

5. Understanding the core principles of indigenous data sovereignty

• Introductory multimedia • video-podcast •

Data Privacy and Consent | Fred Cate | TEDxIndianaUniversity
https://www.youtube.com/watch?v=2iPDpV8ojHA

Arizona State University School of Social Transformation Indigenous Data Sovereignty
https://www.youtube.com/watch?v=TXghvb6lPRI

• Glossary of terms and acronyms • link to Wikipedia when possible •

Datos personales son cualquier información relacionada con una persona identificable. Datos personales

Personal data is any information relating to an identifiable person. Personal data

Data agency is the individual's ability to influence and shape his/her life trajectory as determined by his/her cultural and social contexts. Agency in the digital arena enables an individual to make informed decisions, where his/her own terms and conditions can be recognised and acknowledged at an algorithmic level.
Data agency

Data sovereignty is the idea that data must be subject to the laws and governance structures within the nation in which it is collected. The concept of data sovereignty is closely linked with data security, cloud computing and technological sovereignty. Also, it can be understood as the relation between data and groups of vulnerable or minority groups, which must have agency and voice over how their data is collected, shared and portrayed. Data sovereignty

Data protection is the relationship between the collection and dissemination of data, technology, the public expectation of privacy as well as the legal and political issues surrounding them. It is also known as data privacy. Data protection

GDPR. The General Data Protection Regulation is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA). It also addresses the transfer of personal data outside the EU and EEA areas. The GDPR's primary aim is to give control to individuals over their personal data and to simplify the regulatory environment for international business by unifying the regulation within the EU. GDPR.

• Recommended reading •

1. Matthews, P. (2016). Data literacy conceptions, community capabilities. The Journal of Community Informatics, 12(3). https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3277/4300

2. Kennedy, H, Poell, T., & van Dijck, J. (2015). Data and agency. https://journals.sagepub.com/doi/pdf/10.1177/2053951715621569

3. Drummond, M. (2020). Independent IAM Organizations. IDPro Body of Knowledge, 1(1). https://bok.idpro.org/article/id/32/

4. Schwartz, P. M. (2003). Property, privacy, and personal data. Harv. L. Rev., 117. http://edshare.soton.ac.uk/15267/1/Schwartz-harvard-pdf.pdf

5. The GovLab: Selected Readings on Indigenous Data Sovereignty https://blog.thegovlab.org/post/selected-readings-on-indigenous-data-sovereignty

6. Kukutai & Taylor (Eds.). (2016). Indigenous Data Sovereignty: Toward an agenda. Acton ACT, Australia: ANU Press. http://www.jstor.org/stable/j.ctt1q1crgf

7. Lovett, R., Lee, V., Kukutai, T., Cormack, D., RAINIE, S. C., & Walker, J. (2019). Good data practices for Indigenous data sovereignty and governance. Good Data. Amsterdam: Institute of Network Cultures, 26-36. https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/5b70e9c889858355258ae64a/1534126543958Good+data+practices+for+Indigenous+Data+Sovereignty+and+Governance+submitted.pdf

8. Rainie, S., Kukutai, T., Walter, M., Figueroa-Rodriguez, O., Walker, J., & Axelsson, P. (2019) Issues in Open Data - Indigenous Data Sovereignty. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), The State of Open Data: Histories and Horizons. Cape Town and Ottawa: African Minds and International Development Research Centre. Print version DOI: 10.5281/zenodo.2677801

• Key complementary resources •

1. European Union On-line tool for the security of personal data processing https://www.enisa.europa.eu/risk-level-tool/

2. IEEE Personal Data and Individual Agency https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_personal_data.pdf

3. ICO What are 'controllers' and 'processors'? https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/controllers-and-processors/what-are-controllers-and-processors/

4. Indigenous Data Sovereignty and Governance https://nni.arizona.edu/programs-projects/policy-analysis-research/indigenous-data-sovereignty-and-governance

5. History of Indigenous Data Sovereignty https://www.gida-global.org/history-of-indigenous-data-sovereignty

3 • 1 • Data and personal agency •

> *Because the massive flows of data circulating between devices, institutions, industries and users usher in new and troubling practices of dataveillance that it becomes vital to reflect on whether there are alternative forms of Big Data, forms which enable the less powerful to act with agency in the face of the rise of data power.*
>
> *Kennedy, Poell and van Dijck, 2015*

• Introduction •

According to the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, humans cannot respond on an individual basis to every algorithm tracking their behaviour without technological tools supported by policy allowing them to do so. Individuals may provide consent without fully understanding specific terms and conditions agreements. They are also not equipped to understand how the nuanced use of their data to inform personalised algorithms affects their choices at the risk of eroding their agency. Here, we take agency as the individual's ability to influence and shape his/her life trajectory according to his/her cultural and social contexts. Agency in the digital arena, given conducive social and cultural circumstances, enables an individual to make informed decisions, where his/her own terms and conditions can be recognised and honoured at an algorithmic level.

Fostering agency requires, in particular, enabling the less powerful, the vulnerable and minorities, to have the capabilities they need to challenge unfair decisions and the power dynamics facilitated by data. Students should be able to understand how data collection, processing and usage gives power to some, but not others. This uneven distribution of power creates an imbalance in society that can marginalise those who cannot engage with data effectively, as Atenas, Havemman and Timmermann (2020) assert.

3 • 1.1 • Understanding personal data agency •

For Kennedy, Poell and van Dijck (2015), agency is fundamental when thinking about power data distribution. In the context of datafication, questions about agency have been overshadowed by a focus on oppressive techno-commercial strategies, like data mining (p.2) or even more concerning, what Zuboff (2019) has defined as prediction products that are traded in a new kind of marketplace that she calls behavioural futures markets.

The IEEE recommends governments and organisations provide mechanisms that strengthen individual agency through policies that let individuals create, curate, and control the data associated with their identity. Specifically, they recommend the following.

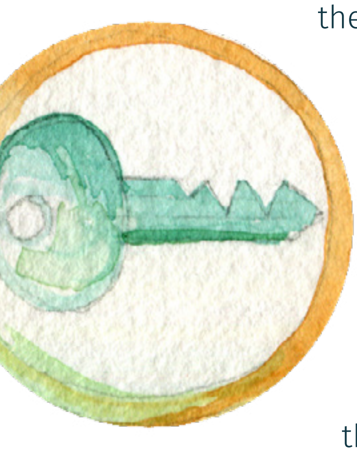
- **Create:** Provide every individual with the means to create and project their own terms and conditions regarding their personal data that can be read and agreed to at a machine-readable level.

- **Curate:** Provide every individual with a personal data or algorithmic agent, which they can curate to represent their terms and conditions in any real, digital, or virtual environment.

- **Control:** Provide every individual access to services allowing them to create a trusted identity to control the safe, specific, and finite exchange of their data.

This is important, as people need to be able to see how their data is being collected by different actors, who are depicting a portrait that can render us subject to all sorts of uses of misuses of automated decisions, thus leading to what is known as the principal–agent problem. According to political science and economics (also known as agency dilemma or the agency problem) scholarship, this occurs when one person or entity (the “agent”) is able to make decisions and/or take actions on behalf of, or that impact upon another person or entity.

The IEEE argues that one of the key challenges is defining how certain uses of data that can affect the individual directly. For example, an individual tube user's travel card can track their movements, so it should be protected from uses that identify or profile that individual so as to be able to make inferences about his/her likes or location generally. Under current business models, it is common for people to consent to the sharing of discrete data, like credit card transaction data, answers to test questions, or how many steps they walk. Once aggregated, these data and the associated insights may lead to complex and sensitive conclusions being drawn about individuals.



The Linking Artificial Intelligence Principles (LAIP) include the concept of contestability, which can be understood as being when an AI system significantly impacts on a person, community, group or environment and that there should be a timely process to allow people to challenge the use or output of the AI system. This principle is aimed at ensuring the provision of efficient, accessible mechanisms that allow people to challenge the use or output of an AI system when it significantly impacts upon a person, community, group or environment. The definition of the threshold for 'significant impact' will depend on the context, impact and application of the AI system in question.

Knowing that redressing for harm is possible when things go wrong, is key to ensuring public trust in AI. Particular attention should be paid to vulnerable persons or groups. There should be sufficient access to the information available to an algorithm, and inferences drawn, to make contestability effective. In the case of decisions significantly affecting rights, there should be an effective system of oversight, which makes appropriate use of human judgment.

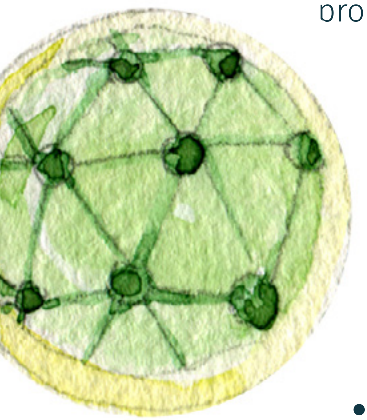
Accordingly, the European Council AI Guidelines propose a framework for Human agency and oversight in which AI systems should support human autonomy and decision-making, as prescribed by the principle of respect for such autonomy. This requires that AI systems acting as enablers to a democratic, flourishing and equitable society by supporting the user's agency and fostering fundamental rights, whilst also allowing for human oversight.

3 • 1.2 • Fostering personal agency •

To enable channels for developing personal agency, we need to foster discussions for students so that they understand that they should be at the centre of their data. They should be encouraged to consider the roadmaps they need to develop to challenge the uses of their data, for people should be able to make informed autonomous decisions regarding AI systems. They should be given the knowledge and tools to comprehend and interact with AI systems to a satisfactory degree and where possible, be enabled to self-assess or challenge the system effectively. A good way to start is discussing the rights of the people regarding AI and then asking whether they can challenge automated decisions. Moreover, there should be discourse on how different countries have different approaches to challenging algorithms. An interesting exercise is to use COVID19 scenarios to identify how such data can be contested.

In his article, Why personal agency matters more than personal data, Searls (2018) states that “The first reason we have far too little agency in the networked world is that we settled, way back in 1995, on a model for websites called client-server, which should have been called calf-cow or slave-master, because we’re always the weaker party: dependent, subordinate, secondary. In defaulted regulatory terms, we clients are mere ‘data subjects’ and only server operators are privileged to be ‘data controllers’, ‘data processors’ or both”. “The second reason agency matters more than data is that nearly the entire market for personal data today is adtech, and adtech is too dysfunctional, too corrupt, too drunk on the data it already has, and absolutely awful at doing what they’ve harvested that data for, which is so machines can guess at what we might want before they shoot ‘relevant’ and ‘interest-based’ ads at our tracked eyeballs”.

Privacy, as a right, needs to be exercised. There are certain abilities needed to do this, one is the ability to understand how data is collected, the terms and conditions and the laws that protect people and allow them to challenge data protection, this ability is of a legal nature. The other ability is more technical and it is related to understanding how different platforms and devices capture data. It is complementary to the prevention and protections of one's data; this ability includes comprehending how encrypted data that has been pseudonymised works and the fact that the process is reversible.



Some of the personal agency data skills that a person needs to develop, can be understood as:

- The ability to understand and give consent for one or more specific purposes;
- Understanding which data is necessary to enter a contract and how one enters a contract;
- To understand the processes needed for fulfilling a legal obligation to protect the vital interests of the user or of another person;
- To understand the necessary process to perform a task carried out in the interest of the public or as contained under the official authority given to the data controller;
- To understand the legitimate interests of the data controller and the rights and freedoms of the user, in particular, those of children.

3 • 1.3 • Data agency and personal rights •

To be able to have personal agency, and to support others in becoming data savvy, it is key that people understand their main data rights, which according to ICO, can be construed as follows.

- **The right to be informed** Refers to the awareness of how organisations must provide users with information about the data processing activities they carry out, normally via a privacy notice/ policy. Users have the right to have this information delivered in a concise, transparent, intelligible, accessible manner, written in clear and plain language (especially if addressed to a child or a vulnerable group of users), and it must be provided free of charge.

- **The right to access** Refers to the right of accessing our own personal data, and to understand information about how the data is being processed. Thus, users need to understand how to request it, and to understand the information data controllers provide them.

- **The right to rectification** It refers to having knowledge of the right to have personal data rectified, if it is biased, inaccurate or incomplete. This right also implies that rectification must be communicated to any third-party recipients involved in the processing of the data, and also to understand how to challenge responses in cases where a request is refused.

- **The right to object** It refers to the ability to understand how to act on the right to object to certain processing activities in relation to personal data. Anyone can object to the processing of their data including research that is of scientific purpose as well as historical research.

- **The right to data portability** It refers to the right of people obtaining (in a machine-readable format) their personal data for the purpose of transferring it from one controller to another, without being prevented from doing so by the data processor. This right only applies to personal data and hence, does not apply to anonymous data.

- **The right to erasure** It refers to the right of having personal data deleted or removed when it is no longer relevant to its original purpose, or the right to withdraw consent or remove such data, if it has been unlawfully processed; requesting that their data be erased and all dissemination ceased. However, the right to erasure can be refused in cases such as where the personal data is processed for archiving purposes in the public interest (for example, scientific research), where it is necessary for legal defence, where the data is necessary to exercise the right of freedom of expression or when it is being processed for health purposes in the public interest.

• Recommended activity •

To challenge automated decisions with the aim of building personal agency, you can work with your students and organise an online group using the Right to Contest activities (it includes downloadable cards).

The Right to Contest objective is to identify AI blind spots that can generate harmful unintended consequences, which arise from our unconscious biases or structural inequalities embedded in society. The right to contest an algorithmic decision can surface inaccuracies and grant agency to the people affected.

After finishing the activities, your students should be encouraged to share their reflections with the rest of the class.

3 • 2 • Data Sovereignty •

La naturaleza multifacética de la soberanía de los datos indígenas da lugar a una amplia gama de cuestiones, desde las dimensiones legales y éticas en torno al almacenamiento, la propiedad, el acceso y el consentimiento de los datos, hasta los derechos de propiedad intelectual y consideraciones prácticas sobre cómo se utilizan los datos en el contexto de la investigación, la política y la práctica

Kukutai and Taylor (2016)

3 • 2.1. Understanding data sovereignty •

Data sovereignty, according to Kukutai & Taylor (2016), is "right to maintain, control, protect and develop their cultural heritage, traditional knowledge and traditional cultural expressions, as well as their right to maintain, control, protect and develop their intellectual property over these". The concept Indigenous Data Sovereignty [ID-SOV] was coined by In 2015 when First Nations scholars and leaders from Australia, Aotearoa/New Zealand, Canada and the United States set up guidelines for the publication of indigenous data (e.g. personal, cultural, historical, land). Hence, non-indigenous people must seek consent in research before publishing any data about indigenous people.

According to IWGIA [in Spanish], the sovereignty of indigenous data is defined as the right of indigenous peoples to own, control, access and possess data that comes from them and that refers to their members, knowledge systems, customs or territories. The sovereignty of indigenous data is grounded in the inherent rights to self-determination and governance over their peoples, territories and resources, as stipulated in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). Thus, it is recognised that indigenous data is a strategic resource and the concept of data sovereignty provides a framework for the ethical use of indigenous information in order to advance the self-determination of the indigenous communities, granting them the right to be decision-makers about how their data is used.

The international approach to the protection of personal data and privacy rights is inadequate for indigenous peoples. Hence, countries need to develop and implement laws, regulations and standards related to the privacy and rights of such people through legal and regulatory approaches co-designed by themselves based on the principles of ID-SOV. ID-SOV can be seen as a driving force in giving indigenous communities the right of self data governance using their values, rights and interests to guide decision-making about how their data is collected, consulted, stored and used. These communities need to be given control of their data through data governance policies and practices as well as through mechanisms and frameworks that reflect their indigenous values.

3 • 2.2 • Principles of indigenous data governance •

The Global Indigenous Data Alliance (GIDA) aims at providing guidance to co-develop frameworks and guidelines for ID-SOV and to disseminate its implementation internationally, through strategic relationships with global bodies and mechanisms. The United Nations Special Rapporteur on the right to privacy has acknowledged the importance of ID-SOV, with the UN Permanent Forum on Indigenous Issues having published some recommendations on Data and Indicators in data disaggregation for indigenous people’s self-determination and development purposes.

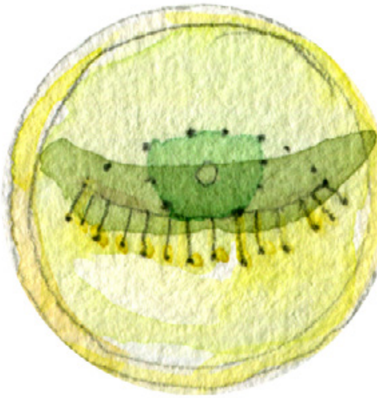
GIDA has published a series of principles for the governance of indigenous data, which includes the right to create value from indigenous data in ways that are grounded in indigenous worldviews and to realise opportunities within the knowledge economy. The four principles can be understood as follows:

1. **Collective benefit:** Data environments should be designed and function in a way that enables indigenous peoples to benefit from the data.

2. **Control authority:** The rights and interests of indigenous peoples over indigenous data must be recognised and their authority to control such data must be empowered.

3. **Responsibility:** Those who work with indigenous data have a responsibility to publicise how that data is used to support self-determination and the collective benefit of indigenous peoples. Accountability requires substantial and openly available evidence of such activities and of the benefits that may be conferred on indigenous peoples.

4. **Ethics:** The rights and well-being of indigenous peoples must be the primary consideration in all phases of the data life cycle and throughout the data environment.



In Australia, the Maiam nayri Wingara Indigneous Data Sovereignty Collective and the Australian Indigenous Governance Institute have developed protocols and principles for Indigenous Data Sovereignty and Indigenous Data Governance, which can be understood as that indigenous peoples have the right to the following.

- Exercise control of the data ecosystem including creation, development, stewardship, analysis, dissemination and infrastructure.

- Data that is contextual and disaggregated (available and accessible at individual, community and First Nations levels).

- Data that is relevant and empowers sustainable self-determination and effective self-governance.

- Data structures that are accountable to indigenous peoples and First Nations.

- Data that is protective and respects our individual and collective interests.

3 • 2.3 • Indigenous peoples and open data •

According to Rainie, Kukutai, Walter, Figueroa-Rodriguez, Walker & Axelsson (2019), when data about indigenous people is opened up without the participation and involvement of indigenous people, it leads to invisibility and bias, while at the same time, provides opportunities for sustainable development.

Thus, for them “Indigenous data sovereignty (IDS) provides a framework for maximising the benefit of open data for Indigenous peoples and other users of Indigenous data and for affecting the stewardship of all data” as “Indigenous nations need data about their citizens, communities, lands, resources, and culture to make informed decisions. Yet, few official statistics agencies, researchers, and data collectors make any meaningful concession to Indigenous rights in relation to Indigenous data. Despite being the rights holders in relation to data about them or for them, Indigenous peoples across nation-states remain peripheral to the channels of power through which consequential decisions about Indigenous statistics are made”.

For Walter, Lovett, Maher, Williamson, Prehn, Bodkin-Andrews & Lee (2020) indigenous people, and communities are normally represented in cases of social disadvantage and they tend to be excluded from the discussions about data collection, governance and use of such data, without acknowledging and recognising indigenous data agency, culture, rights and data needs. This exclusion, in turn, silences their voices during the process of data collection and in so doing marginalises their social, cultural and political stance. Such neo-colonial approaches, more often than not, misappropriate data to the detriment of indigenous people. A way forward to mitigate such neocolonial approaches is to have a pluralistic approach to research and data collection, that is, including the voices of the different stakeholders, including the indigenous community in the open data sector discussion groups. It will also be important to build capacity if needed among the indigenous communities so that they are able to understand how data is collected and then used and how they can exercise their rights when needed.

Recommended activity

To understand how indigenous people are portrayed in data you can ask your students to search for data about indigenous communities and to use the Bias Evaluation Checklist (page 4). Then, get them to write a summary of their findings and encourage them to share their reflections with the rest of the class.

Back to Entendiendo Data: Praxis + Politics

Disclaimer: *unless otherwise stated, all the content produced by the authors for this project is licensed as CC-BY-NC, however, any content produced by third parties such as some of the activities, videos, and other resources may have a different license, so please make sure that you understand each individual licence before using or reusing the content.*

Illustrations, editorial and graphic design