Regulations like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) of 2018 are helping to improve personal data protection. But legal compliance is not enough to mitigate the ethical implications and core challenges to human agency embodied by algorithmically driven behavioral tracking or persuasive computing. The core of the issue is one of parity.

Humans cannot respond on an individual basis to every algorithm tracking their behavior without technological tools supported by policy allowing them to do so. Individuals may provide consent without fully understanding specific terms and conditions agreements. But they are also not equipped to fully recognize how the nuanced use of their data to inform personalized algorithms affects their choices at the risk of eroding their agency.

Here we understand agency as an individual’s ability to influence and shape their life trajectory as determined by their cultural and social contexts. Agency in the digital arena enables an individual to make informed decisions where their own terms and conditions can be recognized and honored at an algorithmic level.

To strengthen individual agency, governments and organizations must test and implement technologies and policies that let individuals create, curate, and control their online agency as associated with their identity. Data transactions should be moderated and case-by-case authorization decisions from the individual as to who can process what personal data for what purpose.

Specifically, we recommend governments and organizations:

- **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 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.

Three sections of this chapter reflect these core ideals regarding human agency.

A fourth section addresses issues surrounding personal data and individual agency relating to children.
Section 1—Create

To retain agency in the algorithmic era, each individual must have the means to create and project their own terms and conditions regarding their personal data. These must be readable and usable by both humans and machines.

**Issue**: What would it mean for a person to have individually controlled terms and conditions for their personal data?

**Background**

Part of providing individually controlled terms and conditions for personal data is to help each person consider what their preferences are regarding their data versus dictating how they need to share it. While questions along these lines are framed in light of a person's privacy, their preferences also reveal larger values for individuals. The ethical issue is whether A/IS act in accordance with these values.

This process of investigating one's values to identify these preferences is a powerful step towards regaining data agency. The point is not only that a person's data are protected, but also that by curating these answers they become educated about how important their information is in the context of how it is shared.

Most individuals also believe controlling their personal data only happens on the sites or social networks to which they belong and have no idea of the consequences of how that data may be used by others in the future. Agreeing to most standard terms and conditions on these sites largely means users consent to give up control of their personal data rather than play a meaningful role in defining and curating its downstream use.

The scope of how long one should or could control the downstream use of their data can be difficult to calculate as consent-based models of personal data have trained users to release rights on any claims for use of their data which are entirely provided to the service, manufacturer, and their partners. However, models like YouTube's [Content ID](https://www.youtube.com/about/services/contentid) provide a form of precedent for thinking about how an individual's data could be technically protected where it is considered as an asset they could control and copyright. Here is language from YouTube's [site about the service](https): “Copyright owners can use a system called Content ID to easily identify and manage their content on YouTube. Videos uploaded to YouTube are scanned against a database of files that have been submitted to us by content owners.” In this sense, the question of how long or how far downstream one’s personal data should be protected takes on the same logic of how long a corporation's intellectual property or copyrights could be protected based on initial legal terms set.
One challenge is how to define use of data that can affect the individual directly, versus use of aggregated data. For example, an individual subway user’s travel card, tracking their individual movements, should be protected from uses that identify or profile that individual to make inferences about his/her likes or location generally. But data provided by a user could be included in an overall travel system’s management database, aggregated into patterns for scheduling and maintenance as long as the individual-level data are deleted. Where users have predetermined via their terms and conditions that they are willing to share their data for these travel systems, they can meaningfully articulate how to share their information.

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. However, once aggregated these data and the associated insights may lead to complex and sensitive conclusions being drawn about individuals. This end use of the individual’s data may not have been part of the initial sharing agreement. This is why models for terms and conditions created for user control typically alert people via onscreen or other warning methods when their predetermined preferences are not being honored.

**Recommendation**

Individuals should be provided tools that produce machine-readable terms and conditions that are dynamic in nature and serve to protect their data and honor their preferences for its use.

**Specifically:**

- Personal data access and consent should be managed by the individual using their curated terms and conditions that provide notification and an opportunity for consent at the time data are exchanged, versus outside actors being able to access personal data without an individual’s awareness or control.
- Terms should be presented in a way that allows a user to easily read, interpret, understand, and choose to engage with any A/IS. Consent should be both conditional and dynamic, where “dynamic” means downstream uses of a person’s data must be explicitly called out, allowing them to cancel a service and potentially rescind or “kill” any data they have shared with a service to date via the use of a “Smart Contract” or specific conditions as described in mutual terms and conditions between two parties at the time of exchange.
- For further information on these issues, please see the following section in regard to algorithmic agents and their application.

**Further Resources**

- IEEE P7012™ - IEEE Standards Project for Machine Readable Personal Privacy Terms. This approved standardization project (currently in development) directly honors the goals laid out in Section One of this document.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 United States License.
Section 2—Curate

To retain agency in the algorithmic era, we must provide every individual with a personal data or algorithmic agent they curate to represent their terms and conditions in any real, digital, or virtual environment. This “agent” would be empowered to act as an individual’s legal proxy in the digital and virtual arena. Oftentimes, the functionality of this agent will be automated, operating along the lines of current ad blockers which do not permit prespecified algorithms to access a user’s data. For other situations that might be unique or new to this agent, a user could specify that notices or updates be sent on a case-by-case basis to determine where there could be a concern.

**Issue:** What would it mean for a person to have an algorithmic agent helping them actively represent and curate their terms and conditions at all times?

**Background**

While it’s essential to create your own terms and conditions to broadcast your preferences, it’s also important to recognize that humans do not operate at an algorithmic speed or level. A significant part of retaining your agency in this
way involves identifying trusted services that can essentially act on your behalf when making decisions about your data.

Part of this logic entails putting you “at the center of your data”. One of the greatest challenges to user agency is that once you give your data away, you do not know how it is being used or by whom. But when all transactions about your data go through your A/IS agent honoring your preferences, you have better opportunities to control how your information is shared.

As an example, with medical data—while it is assumed most would share all their medical data with their spouse most would also not wish to share that same amount of data with their local gym. This is an issue that extends beyond privacy, meaning one’s cultural or individual preferences about what personal information to share, to utility and clarity. This type of sharing also benefits users or organizations on the receiving end of data from these exchanges. For instance, the local gym in the previous example may only need basic heart or general health information and would actually not wish to handle or store sensitive cancer or other personal health data for reasons of liability.

A precedent for this type of patient- or user-centric model comes from Glimpse, a service described by Jordan Crook from TechCrunch in his article, "Apple acquired Glimpse, a personal health data startup": “Glimpse works by letting users pull their own medical info into a single virtual space, with the ability to add documents and pictures to fill out the profile. From there, users can share that data (as a comprehensive picture) to whomever they wish." The fact that Apple acquired the startup points to the potential for the successful business model of user-centric data exchange and putting individuals at the center of their data.

A person’s A/IS agent is a proactive algorithmic tool honoring their terms and conditions in the digital, virtual, and physical worlds. Any public space where a user may not be aware they are under surveillance by facial recognition, biometric, or other tools that could track, store, and utilize their data can now provide overt opportunity for consent via an A/IS agent platform. Even where an individual is not sure they are being tracked, by broadcasting their terms and conditions via digital means, they can demonstrate their preferences in the public arena. Via Bluetooth or similar technologies, individuals could offer their terms and conditions in a ubiquitous and always-on manner. This means even when an individual’s terms and conditions are not honored, people would have the ability to demonstrate their desire not to be tracked which could provide a methodology for the democratic right to protest in a peaceful manner. And where those terms and conditions are recognized meaning technically recognized even if they are not honored one’s opinions could be formally logged via GPS and timestamp data.

The A/IS agent could serve as an educator and negotiator on behalf of its user by suggesting how requested data could be combined with other data that has already been provided, inform the user if data are being used in a way that was not authorized, or make recommendations to the user based on a personal profile. As a negotiator, the agent could broker conditions for sharing data and could include payment to the user as a
term, or even retract consent for the use of data previously authorized, for instance, if a breach of conditions was detected.

**Recommendations**

Algorithmic agents should be developed for individuals to curate and share their personal data. Specifically:

- For purposes of privacy, a person must be able to set up complex permissions that reflect a variety of wishes.
- The agent should help a person foresee and mitigate potential ethical implications of specific machine learning data exchanges.
- A user should be able to override his/her personal agents should he/she decide that the service offered is worth the conditions imposed.
- An agent should enable machine-to-machine processing of information to compare, recommend, and assess offers and services.
- Institutional systems should ensure support for and respect the ability of individuals to bring their own agent to the relationship without constraints that would make some guardians inherently incompatible or subject to censorship.
- Vulnerable parts of the population will need protection in the process of granting access.

**Further Resources**

- IEEE P7006™ – IEEE Standards Project on Personal Data AI Agent Working Group. Designed as a tool to allow any individual to create their own personal “terms and conditions” for their data, the AI Agent will also provide a technological tool for individuals to manage and control their identity in the digital and virtual world.
- Tools allowing an individual to create a form of an algorithmic guardian are often labeled as PIMS, or Personal Information Management Services. Nesta in the United Kingdom was one of the funders of early research about PIMS conducted by CtrlShift.
Section 3—Control

To retain agency in the algorithmic era, we must provide every individual access to services allowing them to create a trusted identity to control the safe, specific, and finite exchange of their data.

**Issue:** How can we increase agency by providing individuals access to services allowing them to create a trusted identity to control the safe, specific, and finite exchange of their data?

**Background**

Pervasive behavior-tracking adversely affects human agency by recognizing our identity in every action we take on and offline. This is why identity as it relates to individual data is emerging at the forefront of the risks and opportunities related to use of personal information for A/IS. Across the identity landscape there is increasing tension between the requirement for federated identities versus a range of identities. In federated identities, all data are linked to a natural and identified person. When one has a range of identities, or personas, these can be context specific and determined by the use case. New movements, such as “Self-Sovereign Identity”—defined as the right of a person to determine his or her own identity—are emerging alongside legal identities, e.g., those issued by governments, banks, and regulatory authorities, to help put individuals at the center of their data in the algorithmic age.

Personas, identities that act as proxies, and pseudonymity are also critical requirements for privacy management and agency. These help individuals select an identity that is appropriate for the context they are in or wish to join. In these settings, trust transactions can still be enabled without giving up the “root” identity of the user. For example, it is possible to validate that a user is over eighteen or is eligible for a service.

Attribute verification will play a significant role in enabling individuals to select the identity that provides access without compromising agency. This type of access is especially important in dealing with the myriad of algorithms interacting with narrow segments of our identity data. In these situations, individuals typically are not aware of the context for how their data will be used.

**Recommendation**

Individuals should have access to trusted identity verification services to validate, prove, and support the context-specific use of their identity.

**Further Resources**

Section 4—Children’s Data Issues

While the focus of this chapter is to provide all individuals with agency regarding their personal data, some sectors of society have little or no control. For some elderly individuals or the mentally ill, it is because they have been found not to have “mental capacity”, and for prisoners in the criminal justice system, society has taken control away as punishment. In the case of children, this is because they are considered human beings in development with evolving capacities.

We examine the issues of children as an example case and recommend either regulation or a technical architecture that provides a veil and buffer from harm until a child is at an age where they can claim personal responsibility for their decisions.

In many parts of the world, children are viewed by the law as being primarily charges of their parents who make choices on their behalf. In Europe, however, the state has a role in ensuring the “best interests of the child”. In schools, the two interests operate side-by-side, with parents being given some control over their child’s education but with many decisions being made by the schools.

Many of the issues described above concern choices around personal data and the future impacts of how the data are gathered and shared. Children are at the forefront of technological developments with future educational and recreational technology gathering data from them all day at school and intelligent toys throughout their time at home.

As children post, click, search, and share information, their data are linked to various profiles, grouped into segmented audiences, and fed into machine learning algorithms. Some of these may be designed to target campaigns that increase sales, influence sentiment, encourage online games, impact social networks, or influence religious and political views. Data fed into algorithmic advertising is not only gathered from children’s online actions but also from their devices. An example of device data is browser fingerprinting. It includes a set of data about a child’s browser or operating system. Fingerprinting vastly increases privacy risks because it is used to link to an individual.

Increasingly, children’s beliefs and social norms are established by what they see and experience online. Their actions reflect what they believe is possible and expected. The report, "Digital Deceit: Technologies Behind Precision Propaganda on the Internet", explains how companies collect, process, and then monetize personal preferences, socioeconomic status, fears, political and religious beliefs, location, and patterns of internet use.

Companies, governments, political parties, and philosophical and religious organizations use data available about students and children to influence how they spend their time, money, and the people or institutions they trust and with whom they spend time and build relationships.

Many aspects of a child’s life can be digitized. Their behavioral, device, and network data are combined and used by machine learning
algorithms to determine the information and content that best achieve the educational goals of the schools and the economic goals of the advertisers and platform companies.

**Issue: Mass personalization of instruction**

**Background**

The mass personalization of education offers better education for all at very low cost through A/IS-enabled computer-based instruction that promises to free up teachers to work with kids individually to pursue their passions. These applications will rely on the continuous gathering of personal data regarding mood, thought processes, private stories, physiological data, and more. The data will be used to construct a computational model of each child’s interests, understanding, strengths, and weaknesses. The model provides an intimate understanding of how they think, what they understand, how they process information, or react to new information; all of which can be used to drive instructional content and feedback.

Sharing of this data between classes, enabling it to follow students through their schooling, will make the models more effective and beneficial to children, but it also exposes children and their families to social control. If performance data are correlated with social data on a family, it could be used by social authorities in decision-making about the family. For example, since 2015-2018, well-being digital tests were performed in schools in Denmark. Children were asked about everything from bullying, loneliness, and stomachaches. Recently it was disclosed that although the collected data was presented as anonymous, they were not. Data were stored with social security numbers, correlated with other test data, and even used in case management by some Danish municipalities. Commercial profiling and correlation of different sets of personal data may further affect these children in future job or educational situations.

**Recommendation**

Educational data offer a unique opportunity to model individuals’ thought processes and could be used to predict or change individuals’ behavior in many situations. Governments and organizations should classify educational data as being sensitive and implement special protective standards.

Children’s data should be held in “escrow” and not used for any commercial purposes until a child reaches the age of majority and is able to authorize use as they choose.

**Further Resources**


### Issue: Technology choice-making in schools

**Background**

Children, as minors, have no standing to give or deny consent, or to control the use of their personal data. Parents only have limited choices in what are often school-wide implementations of educational technology. Examples include the use of Google applications, face recognition in security systems, and computer driven instruction as described above. In many cases, parents’ only choice would be to send their children to a different school, but that choice is seldom available.

How should schools make these choices? How much input should parents have? Should parents be able to demand technology-free teaching?

There are many gaps in current student data regulation. In June 2018, CLIP, The Center on Law and Information Policy at Fordham Law School published, “Transparency and the Marketplace for Student Data”6. This study concluded that “student lists are commercially available for purchase on the basis of ethnicity, affluence, religion, lifestyle, awkwardness, and even a perceived or predicted need for family planning services”. Fordham found that the data market is becoming one of the largest and most profitable marketplaces in the United States.

Data brokers have databases that store billions of data elements on nearly every United States consumer. However, information from students in the pursuit of an education should not be exploited and commercialized without restraint.

Fordham researchers found at least 14 data brokers who advertise the sale of student information. One sold lists of students as young as two years old. Another sold lists of student profiles on the basis of ethnicity, religion, economic factors, and even gawkiness.

**Recommendation**

Local and national educational authorities must work to develop policies surrounding students’ personal data with all stakeholders: administrators, teachers, technology providers, students, and parents in order to balance the best educational interests of each child with the best practices to ensure safety of their personal data. Such efforts will raise awareness among all stakeholders of the promise and the compromises inherent in new educational technologies.
Personal Data and Individual Agency

Further Resources

- Common Sense Media privacy evaluation project: [https://www.commonsense.org/education/privacy](https://www.commonsense.org/education/privacy)


Issue: Intelligent toys

Background

Children will not only be exposed to A/IS at school but also at home, while they play and while they sleep. Toys are already being sold that offer interactive, intelligent opportunities for play. Many of them collect video and audio data which is stored on company servers and either is or could be mined for profiling or marketing data.

There is currently little regulatory oversight. In the United States COPPA offers some protection for the data of children under 13. Germany has outlawed such toys using legislation banning spying equipment enacted in 1981. Corporate A/IS are being embodied in toys and given to children to play with, to talk to, tell stories to, and to explore all the personal development issues that we learn about in private play as children.

Recommendations

Child data should be held in “escrow” and not used for any commercial purposes until a child reaches the age of majority and is able to authorize use as they choose.

Governments and organizations need to educate and inform parents of the mechanisms of A/IS and data collection in toys and the possible impact on children in the future.

Further Resources


Personal Data and Individual Agency


Thanks to the Contributors

We wish to acknowledge all of the people who contributed to this chapter.

The Personal Data and Individual Agency Committee

• Katryna Dow (Co-Chair) – CEO & Founder at Meeco
• John C. Havens (Co-Chair) – Executive Director, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems; Executive Director, The Council on Extended Intelligence; Author, Heartificial Intelligence: Embracing Our Humanity to Maximize Machines
• Mads Schaarup Andersen – Senior Usable Security Expert in the Alexandra Institute’s Security Lab

• Ajay Bawa – Technology Innovation Lead, Avanade Inc.
• Ariel H. Brio – Privacy and Data Counsel at Sony Interactive Entertainment
• Walter Burrough – Co-Founder, Augmented Choice; PhD Candidate (Computer Science) – Serious Games Institute
• Danny W. Devriendt – Managing director of Mediabrands Dynamic (IPG) in Brussels, and the CEO of the Eye of Horus, a global thinktank for communication-technology related topics
• Dr. D. Michael Franklin – Assistant Professor, Kennesaw State University, Marietta Campus, Marietta, GA
• Jean-Gabriel Ganascia – Professor, University Pierre et Marie Curie; LIP6 Laboratory ACASA Group Leader
Personal Data and Individual Agency

- **Bryant Joseph Gilot, MD CM DPhil MSc** – Center for Personalised Medicine, University of Tuebingen Medical Center, Germany & Chief Medical Officer, Blockchain Health Co., San Francisco
- **David Goldstein** – Seton Hall University
- **Adrian Gropper, M.D.** – CTO, Patient Privacy Rights Foundation; HIE of One Project
- **Marsali S. Hancock** – Chair, IEEE Standards for Child and Student Data governance, CEO and Co-Foundation EP3 Foundation
- **Gry Hasselbalch** – Founder DataEthics, Author, *Data Ethics - The New Competitive Advantage*
- **Yanqing Hong** – Graduate, University of Utrecht Researcher at Tsinghua University
- **Professor Meg Leta Jones** – Assistant Professor in the Communication, Culture & Technology program at Georgetown University
- **Mahsa Kiani** – Chair of Student Activities, IEEE Canada; Vice Editor, IEEE Canada Newsletter (ICN); PhD Candidate, Faculty of Computer Science, University of New Brunswick
- **Brenda Leong** – Senior Counsel, Director of Operations, The Future of Privacy Forum
- **Emma Lindley** – Founder, Innovate Identity
- **Ewa Luger** – Chancellor’s Fellow at the University of Edinburgh, within the Design Informatics Group
- **Sean Martin McDonald** – CEO of FrontlineSMS, Fellow at Stanford’s Digital Civil Society Lab, Principal at Digital Public
- **Hiroshi Nakagawa** – Professor, The University of Tokyo, and AI in Society Research Group Director at RIKEN Center for Advanced Intelligence Project (AIP)
- **Sofia C. Olhede** – Professor of Statistics and an Honorary Professor of Computer Science at University College London, London, U.K.; Member of the Programme Committee of the International Centre for Mathematical Sciences.
- **Ugo Pagallo** – University of Turin Law School; Center for Transnational Legal Studies, London; NEXA Center for Internet & Society, Politecnico of Turin
- **Dr. Juuso Parkkinen** – Senior Data Scientist, Nightingale Health; Programme Team Member, MyData 2017 conference
- **Eleonore Pauwels** – Research Fellow on AI and Emerging Cybertechnologies, United Nations University (NY) and Director of the AI Lab, Woodrow Wilson International Center for Scholars (DC)
- **Dr. Deborah C. Peel** – Founder, Patient Privacy Rights & Creator, the International Summits on the Future of Health Privacy
- **Walter Pienciak** – Principal Architect, Advanced Cognitive Architectures, Ltd.
- **Professor Serena Quattrocolo** – University of Turin Law School
- **Carolyn Robson** – Group Data Privacy Manager at Etihad Aviation Group
- **Gilad Rosner** – Internet of Things Privacy Forum; Horizon Digital Economy Research Institute, UK; UC Berkeley Information School
Personal Data and Individual Agency

- **Prof. Dr.-Ing. Ahmad-Reza Sadeghi** – Director System Security Lab, Technische Universität Darmstadt / Director Intel Collaborative Research Institute for Secure Computing
- **Rose Shuman** – Partner at BrightFront Group & Founder, Question Box
- **Dr. Zoltán Szlávik** – Lead/Researcher, IBM Center for Advanced Studies Benelux
- **Udbhav Tiwari** – Centre for Internet and Society, India

For a full listing of all IEEE Global Initiative Members, visit standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ec_bios.pdf.

For information on disclaimers associated with EAD1e, see How the Document Was Prepared.

Endnotes


5. Case described in Danish here https://dataethics.eu/trivsel-enhver-pris/
