# APPENDIX A: UNDERSTANDING DATA PRIVACY

In order to most effectively engage and communicate with stakeholders about privacy, IDS should be familiar with the basics of modern privacy and data protection. Understanding common privacy principles, risks, and resources will help focus your efforts to gather community input and create appropriate privacy safeguards.

## A.1: Privacy basics

Data privacy is...

- **A fundamental right.**[10] Individual privacy rights are recognized in the U.S. Constitution, the Declaration of Human Rights, and in over 80 countries around the world. Privacy rights also provide the foundation for other important rights, including self-determination and free expression.
- **Dynamic.** Some of the most common aspects of data privacy include:
  - Control over personal information flows;
  - Freedom of thought and exploration; and
  - Protection of one's dignity and reputation.
- **Subjective.** Each person has unique privacy preferences and expectations—what feels invasive or creepy to one person may be innovative or cool to another. These preferences and expectations are influenced by many factors, including a person's familiarity with data systems, experiences of marginalization, cultural background, and trust in data-holding organizations.[11]
- **Contextual.** Whether it is appropriate to use or share personal data in a particular manner depends on ever-evolving social and ethical norms, as well as legal frameworks. In order to understand your community's norms about integrating data, it is essential to communicate and engage directly with members of your community.
- **Power.** The more information that one person or organization has about another, typically, the more that party may influence or exert power over the other. Privacy protections help individuals and communities maintain their autonomy and freedom when their information is used by governments and other organizations.

Data privacy is *not*...

- **Secrecy.** Seeking to protect privacy does not mean preventing all others from learning information about you. On the contrary, privacy is about creating conditions in which individuals will trust others enough to share their personal information.
- **Security.** Although data privacy and data security are closely related, a perfectly secure data system may still violate individual privacy if authorized users acting within an organization or system's normal capabilities use personal data in unexpected or inappropriate ways.[12]

Although privacy definitions may seem abstract and subjective, privacy risks can be significant and concrete. Failing to adequately safeguard data or using it inappropriately can have serious and lasting consequences for individuals.

Common privacy risks for individuals include:[13]

- **Financial** risks, such as identity theft or fraud;
- **Physical risks,** such as stalking or burglary;
- **Reputational risks,** such as embarrassing rumors or damaging photos; and
- **Dignitary risks,** such as a loss of autonomy or opportunity when a person is profiled or discriminated against by an automated decision-making system.

Nevertheless, privacy risks may not always be immediately apparent. Many times, privacy risks for individuals are...

- **Incremental.** As datasets grow and are combined over time, so does the likelihood of a data breach, a successful re-identification attack (singling out individuals in seemingly non-personal data), or a discriminatory impact on vulnerable, historically marginalized, and/or over-surveilled communities.
- **Unequal.** Privacy risks may also accrue unevenly throughout society. If not addressed in advance, some community members may reap the benefits of data-driven governance, while others bear all of the privacy risk burden.
- **Non-obvious.** Certain privacy risks are more impactful or more likely to occur for particular groups, and can be overlooked by program

designers who have not specifically incorporated those individuals' inputs. For example, publishing the contact information of everyone who attended a meeting will have different implications for an elected official than a domestic violence survivor.

- **Intrusive.** Privacy is closely tied to feelings about self-control and autonomy, and its real—or perceived—loss can leave people feeling vulnerable, exposed, and out of control of their own lives. In these situations, individual and community behavior can be chilled, relationships harmed, and trust lost.

Because privacy risks can be so varied, it is critical to engage diverse stakeholders in discussions and decision-making about integrating personal data. Incorporating non-traditional voices within your stakeholder groups will strengthen your ability to foresee and address privacy or equity externalities arising from the IDS' work.

## A.2: Integrated data privacy

Privacy risks are not hypothetical, and neither are public responses to them. Failure to protect individuals' privacy—including miscommunications or silence by organizations about how personal data will be used and protected—can lead to lasting public mistrust,[14] internal protests,[15] project collapse,[16] and even legislative backlash.[17]

Common risks for organizations (including IDS) that fail to handle privacy correctly, or are *perceived* to fail at privacy, include:

- **Financial** risks, such as lawsuits or statutory damages, or the withdrawal of funding;
- **Reputational** risks, such as loss of public trust and support in the IDS; and
- **Regulatory,** such as new legislative restrictions on administrative data use and sharing.

IDS must be mindful that, as representatives of state and local government, your use of individuals' personal data can be particularly fraught. IDS face several unique hurdles to earning public trust and social license to use their community's personal data, including:

- **History.** Over the course of history, both governments and researchers have mistreated and misused personal data.[18] IDS should be aware that those scars have lingered, and appreciate that some individuals and communities have valid reasons to be reluctant about data sharing.
- **Big Brother looms large.** When data is collected and used by government institutions, privacy risks—and fears—can become amplified. To many communities, frequent data collection by government agencies, even for beneficial purposes, can feel indistinguishable from Big Brother-type surveillance. Historically marginalized populations, such as people of color and those living in poverty, in particular may be the target of multiple, concurrent data collection efforts by local, state, and federal agencies.
- **Consent.** Most privacy laws require organizations to get an individuals' consent before using their personal information, particularly when the data is sensitive. The nature of most IDS activities, however, means that they are secondary uses of administrative data and such consent isn't feasible. While IDS activities *are* specifically permitted under those same laws, many people may be unfamiliar with those exceptions or may simply expect an opportunity to consent by default.
- **Data-driven inequalities.** There is a growing public, private, and academic conversation about how data-driven tools may reflect or reinforce discrimination and bias, even inadvertently.[19] The use of administrative data by algorithmic systems, whether by IDS or other organizations, raise serious ethical questions and may color public perceptions of *other* data uses.
- **Trust.** The perception that personal data will be used in unexpected ways or will not be kept private and secure can undermine individuals' and communities' trust in government. At its extreme, individuals who are afraid of how data about them could be used may even provide false information or forgo government services.[20]

Meaningful stakeholder engagement and communication, as well as strong privacy protections, can help put power back in the hands of individuals and communities. Protecting privacy is critical to respecting individuals' rights and maintaining individuals' trust in government.

Some uses of personal data will carry more inherent privacy risks than others. For IDS, these include...

| LOWER RISK | HIGHER RISK |
|---|---|
| Non-sensitive data (e.g., demographic or contact information) | Sensitive data (e.g., health, financial, criminal justice, location, or education data)[21] |
| Aggregated data | Individual records |
| Data about groups | Data about individuals |
| Cross-sectional research | Longitudinal research |
| Evaluating outcomes | Predicting outcomes |
| Policy analysis and research | Case management, enforcement, adjudication |

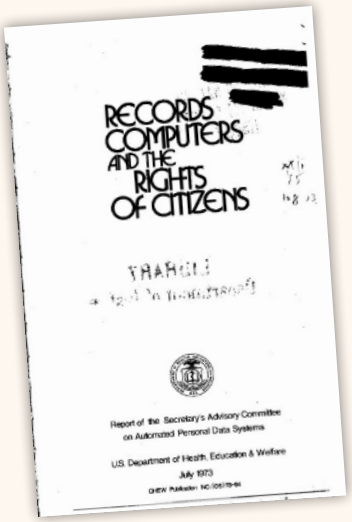
Make sure your IDS is prepared to differentiate between these distinct types of data and distinct data uses in conversations about privacy risks and potential benefits.

The goals of IDS serve important social, economic, and democratic functions; however, if citizens do not trust that their data will be protected or do not see the benefits of IDS use for research and evaluation, they could begin to fear administrative data and government services as tools of surveillance, rather than tools for change.

## A.3: Privacy Fair Information Practice Principles (FIPPs)

In order to effectively address privacy risks, privacy professionals and policymakers look to the Fair Information Practice Principles (FIPPs). The FIPPs serve as the common language of privacy and are the basis of all privacy laws around the world. Originating in the 1970s and 80s, the FIPPs consist of eight core principles[22]:

- **Collection Limitation Principle:** There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.
- **Data Quality Principle:** Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.
- **Purpose Specification Principle:** The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.




- **Use Limitation Principle:** Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with the purpose specification principle except with the consent of the data subject or by the authority of law.
- **Security Safeguards Principle:** Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.
- **Openness Principle:** There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.
- **Individual Participation Principle:** An individual should have the right:
  - to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him; to have communicated to him, data relating to him within a reasonable time; at a charge, if any, that is not excessive; in a reasonable manner; and in a form that is readily intelligible to him;
  - to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and
  - to challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed or amended.
- **Accountability Principle:** A data controller should be accountable for complying with measures which give effect to the principles stated above.

For many IDS, there may be specific laws or policies that dictate how personal data should be handled and protected for certain activities. But there may also be times that the laws and policies do not clearly capture IDS activities, or where an IDS wants to go above and beyond basic compliance. In these situations, IDS can look to the FIPPs for guidance.

In general, IDS activities lend themselves more strongly to some FIPPs than others. The consent and use/purpose limitation principles in particular seem to conflict with the goals of IDS. However, other FIPPs directly support a commitment to privacy-protective integrated data activities, including the principles of accountability, openness, security, and, in some circumstances, individual participation.

Importantly, IDS should remember that the FIPPs are *principles* for privacy protections, not absolute laws. They can and should be adopted flexibly, to maximize the benefits of integrated data while minimizing the risks to individual privacy. For example, while IDS may find that gathering individual consent is infeasible, they may place a stronger emphasis on other FIPPs, such accountability and openness, through robust stakeholder engagement processes.

## A.4: Privacy tools and resources

Ultimately, effective engagement and communication on privacy issues must be grounded in strong data policies and practices. While this section is not intended to provide the components of a comprehensive IDS privacy program, it will help IDS stakeholders and their communities explore key privacy tools and resources. IDS should consult their legal counsel and organizational leadership, as well as appropriate stakeholders, in developing and implementing relevant privacy safeguards.

Some common privacy tools and resources to consider include:

**Privacy Control Catalogues** describe specific technical and administrative safeguards that can be used to protect and manage data flows.

- National Institute of Standards and Technology (NIST), *SP 800-53 Rev 4: Security and Privacy Controls for Federal Information Systems and Organizations*, *Appendix J: Privacy Control Catalog* (2013), https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-53r4.pdf.
- Centers for Medicare & Medicaid Services, *Volume III: Catalog of Minimum Acceptable Risk Security and Privacy Controls for Exchanges, Version 2.0* (2015), https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/3-MARS-E-v2-0-Catalog-of-Security-and-Privacy-Controls-11102015.pdf.

**Privacy Impact Assessments** are analyses of how personally identifiable information is collected, used, shared, and maintained by an organization, and is typically used to identify specific privacy risks.

- NIST, *NISTIR 8062: Privacy Risk Management for Federal Information Systems* (2017), http://csrc.nist.gov/publications/drafts/nistir-8062/nistir_8062_draft.pdf.
- Bureau of Justice Assistance, U.S. Department of Justice, *Guide to Conducting Privacy Impact Assessments: for State, Local, and Tribal Justice Entities* (2012), https://it.ojp.gov/documents/d/Guide%20to%20Conducting%20Privacy%20Impact%20Assessments_compliant.pdf.
- Information Commissioner's Office (UK), *Data Protection Impact Assessments* (Aug. 2018), https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/.
- National Commission on Informatics & Liberty (Commisson nationale de l'linformatique et des libertés or CNIL), *Privacy Impact Assessment (PIA): Knowledge Bases* (Feb. 2018), https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-3-en-knowledgebases.pdf.

**Privacy programs** ensure that responsibility is established, accountability is maintained, and resources are allocated within an organization to successfully oversee, govern, and use individuals' data.

- Office of Management and Budget (OMB), 81 FR 49689, *OMB Circular No. A-130: Managing Information as a Strategic Resource* (July 2016), https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf.
- Organization for Economic Cooperation and Development (OECD), *OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, C(80)58/FINAL (July 11, 2013), https://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf.
- Daniel Solove and Woodrow Hartzog, *The Federal Trade Commission and the New Common Law of Privacy*, 114 Colum. L. Rev. 583 (2014), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2312913.
- European Commission, *Guidelines on Data Protection Officers ('DPOs')* (Dec. 13, 2016), http://ec.europa.eu/newsroom/document.cfm?doc_id=43823.
- Privacy Technical Assistance Center (PTAC), *Data Governance Checklist* (June 2015), https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Data%20Governance%20Checklist_0.pdf.

**De-Identification** is the process of removing or perturbing identifiable data elements such that individuals can no longer be identified, singled out, or linked to other attributes through their data.

- NIST, *NISTIR 8053: De-Identification of Personal Information 2* (2015), http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR .8053.pdf.
- NIST, *SP 800-188 (2nd Draft): De-Identifying Government Datasets* (Dec. 15, 2016), http://csrc.nist.gov/publications/drafts/800-188/sp800_188_draft.pdf.
- National Academies of Sciences, Engineering and Medicine, *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy* (2007), https://www.nap.edu/catalog/24652/innovations-in-federal-statistics-combining-data-sources-while-protecting-privacy.

- National Academies of Sciences, Engineering, and Medicine, *Federal Statistics, Multiple Data Sources, And Privacy Protection: Next Steps* (2017), https://www.nap.edu/catalog/24893/federal-statistics-multiple-data-sources-and-privacy-protection-next-steps.
- Micah Altman et al., *Towards a Modern Approach to Privacy-Aware Government Data Releases*, 30 Berkley Tech. L. J. 1967 (2015), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2779266.
- Ira Rubinstein and Woodrow Hartzog, *Anonymization and Risk*, 91 Wash. L. Rev. 703 (2016), http://lsr.nellco.org/cgi/viewcontent.cgi?article=1534&context=nyu_plltwp.
- Future of Privacy Forum, *A Visual Guide to Practical Data De-Identification*(Apr. 25, 2016), https://fpf.org/2016/04/25/a-visual-guide-to-practical-data-de-identification.
- Future of Privacy Forum, *Open Data Privacy Risk Assessment for the City of Seattle*, *Appendix C* (Jan. 30, 2018), https://fpf.org/wp-content/uploads/2018/01/FPF-Open-Data-Risk-Assessment-for-City-of-Seattle.pdf.


**Data Ethics Frameworks** are tools and considerations for helping evaluate whether using data use is unethical and/or whether the benefits and risks are unbalanced.


- Data for Democracy, BrightHive, and Bloomberg, *Community Principles on Ethical Data Practices* (Sept 2017), https://datapractices.org/community-principles-on-ethical-data-sharing.
- Markkula Center for Applied Ethics, *A Framework for Ethical Decision-Making,* Santa Clara University (Aug. 1, 2015), https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/a-framework-for-ethical-decision-making/.
- Department for Digital, Culture, Media & Sport (UK), *Data Ethics Workbook* (June 13, 2018), https://www.gov.uk/government/publications/data-ethics-workbook/data-ethics-workbook.
- Open Data Institute (ODI), *The Data Ethics Canvas* (Aug. 5, 2017), https://theodi.org/article/data-ethics-canvas/.
- Future of Privacy Forum, *Benefit-Risk Analysis for Big Data Projects* (Sept. 2014), https://fpf.org/wp-content/uploads/FPF_DataBenefitAnalysis_FINAL.pdf.
- Omer Tene and Jules Polonetsky, *Beyond IRBs: Ethical Guidelines for Data Research*, 732 Wash. & Lee L. Rev. Online 458 (2016), https://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=1044&context=wlulr-online.