Data Ethics Framework

September 2020

Contents

1 Overview

Decisions made with data touch every aspect of American life. The Federal Government uses data to solve problems, develop and deliver services to citizens, defend and secure the nation, and support economic growth. Data's benefits and risks are amplified by the expanding capabilities of digital networks, IT systems, algorithms, and computational methods that enable data to be easily collected, combined, manipulated, and shared.

The Federal Data Strategy, delivered in December 2019, recognized the importance of ethics in its founding Principles. When the Federal Data Strategy team created the 2020 Action Plan, they specifically tasked the General Services Administration (GSA) with developing a Data Ethics Framework (Framework) in Action 14 to help agency employees, managers, and leaders make ethical decisions as they acquire, manage, and use data.

To achieve the goal of developing a useful Framework for the Federal Government, GSA formed an Inter-Agency Team (IAT) comprised of 14 government leaders from different agencies, with expertise in statistics, public policy, evidence-based decision making, privacy, and analytics. GSA also received input on the Framework from the Chief Data Officer (CDO) Council, Interagency Committee on Standards Policy (ICSP), and the Federal Privacy Council (FPC). The resulting Framework is intended to be a "living" resource and to be regularly updated by the CDO Council and ICSP.

The Framework incorporates the input and terminology from stakeholders representing many domains, and who use different types of data in different ways. The developers of the Framework recognize that some terms may be used differently, depending on the context, type of data being used, and stage in the data lifecycle. The Framework applies to all data types and data uses.

The Framework consists of four parts:

- **About the Data Ethics Framework** outlines the intended purpose and audience of this document
- **Data Ethics Defined** explores the meaning of the term "data ethics," as background to the Tenets provided in the following section
- **Data Ethics Tenets** provides seven Tenets, or high-level principles, for using data ethically within the Federal Government
- **Data Ethics Tenets in Action** describes the benefits of data ethics and contains use cases demonstrating how the Tenets can guide data activities within federal agencies and federally sponsored programs

2 About the Data Ethics Framework

2.1 Purpose

The Framework's purpose is to guide ethical decision making by federal employees who collect, manage, and use data in order to support their agency's mission. The Framework does not include requirements or mandates of its own but provides guidance in the form of foundational principles, called Tenets, to encourage ethical decision making at all levels of the Federal government.

2.2 Audience

The Framework is for anyone in the Federal Government who works with or leads work involving data, which includes all employees, contractors, researchers, and other partners who work on behalf of the government.

In particular, the Framework is relevant to those who work with data during any stage of the data lifecycle (Figure 1), including collecting, processing, disseminating, using, or storing and disposing of data. That includes:

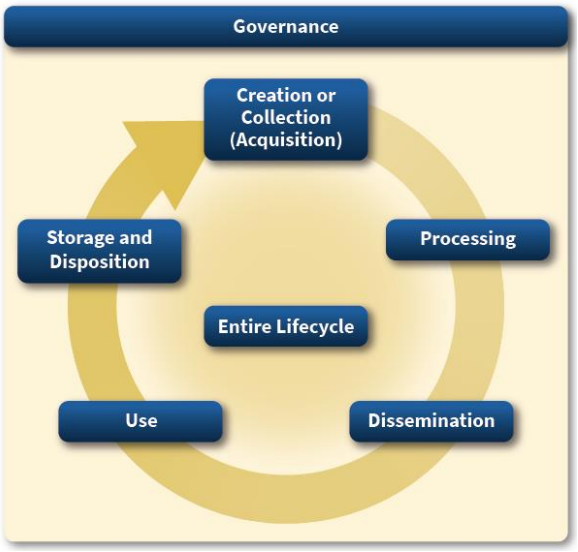

Figure 1: Data Lifecycle adopted from OMB Circular A-130, Managing Information as a Strategic

- Organizational leaders, including CDOs, heads of agencies, senior executives, and supervisors
- Data practitioners, such as statisticians, data analysts, database professionals, and data scientists
- Operational employees that collect, use, manage, and report data for regular program operations
- Policymakers and those advising decision makers
- Data stewards, both those who manage data for programmatic purposes and those who manage administrative data, such as human resource employees
- Public relations officials, communications employees, and agency representatives who are presenting information and data to the public
- Data consumers, such as other agencies, communities, or the public

3 Data Ethics Defined

3.1 Data Ethics Definition

Data ethics are the norms of behavior that promote appropriate judgments and accountability when collecting, managing, or using data, with the goals of protecting civil liberties, minimizing risks to individuals and society, and maximizing the public good.

Remaining a leader in data ethics requires individuals, federal agencies, and cross-agency communities to embrace ongoing discussion and learning. Instead of looking at issues from a single perspective, ethical decision making is best achieved by taking a holistic approach and widening the context to weigh the greater implications of data use.

3.2 Application of Data Ethics

The Data Ethics Tenets apply to all data types and data uses. It is understood that the same dataset may be used at different times for different purposes. No matter the data type or use, federal employees should ensure the protection of privacy (i.e., state of being free from unwarranted intrusion into the private life of individuals), confidentiality (i.e., free from inappropriate access and use), civil rights, and civil liberties during data activities.

In addition, this Framework lives in a context of other federal ethics guidance and requirements. For example, the Privacy Act of 1974 established the U.S. Privacy Protection Study Commission (PPSC), charging the group with providing legislative recommendations to help protect the privacy of individuals while meeting the legitimate needs of government and society for information. The PPSC realized that most Americans treasure their personal privacy, but they are also willing to give information about themselves if it drives some concrete benefit (e.g., social security).

The Federal Government recognizes that work with some data, such as operational datasets, requires less scrutiny and is inherently low risk. There are also scenarios where work with certain data, such as information protected by the Privacy Act, requires more scrutiny and a full assessment of potential impacts. Each community or agency is ultimately responsible for applying this Framework appropriately based on the context and level of risk.

To support an infrastructure that protects individual privacy while allowing for the use of information to facilitate the delivery of services, the Federal Government distinguishes between two fundamental uses of data:

- Uses that affect the rights, privileges, and benefits of an individual
- Uses that do not affect an individual's rights, as the data are used only to create aggregated results

This distinction enables Federal Government to ensure data are effectively protected while also providing for the ability to better understand the effectiveness of government services.

4 Data Ethics Tenets

The Federal Data Ethics Tenets are intended to help Federal employees make decisions ethically and promote accountability throughout the data lifecycle. Regardless of data type or use, those working with data in the public sector should have a foundational understanding of the Data Ethics Tenets, and leaders should continuously strive to support a data ethics-driven culture and lead by example.

Each Tenet includes the following:

- **Recommendations for Federal Leaders and Employees** who work with data
- **Legal Authorities** that govern information and data use in the Federal Government
- **Additional Resources** to help inform responsible approaches to data use

The legal authorities and additional resources provided are not exhaustive but are recognized as relevant to federal data ethics.

4.1 Be Aware of and Uphold Applicable Statutes, Regulations, Professional Practices, and Ethical Standards

Data leaders and professionals should adhere to all applicable legal authorities, as ethics are reflected and reinforced in existing laws. At the same time, legal authorities are developed to address historic situations and issues and may not keep pace with the evolving world of data and technology. Therefore, agency leaders are encouraged to maintain up-to-date, comprehensive ethical standards regarding data use and staff are responsible for learning and applying agency guidance. In addition, if a person works in an area with recognized professional ethical codes of conduct, they should be aware of those standards and strive to uphold them.

Recommendations for Federal Leaders:

- Identify and clearly communicate the legal authorities, professional codes of conduct, and ethical standards that apply to their organization
- Support agency-level ethical standards that include components covering data ethics
- Identify and clearly communicate how different types of data within the organization should be handled
- Recognize the diverse roles of employees working with data throughout the data lifecycle and provide clear guidance and instruction based on the employee's data role and level within the organization
- Provide training and learning opportunities to increase employee knowledge in the areas of applicable statutes, regulations, professional practices, and ethical standards
- Implement mechanisms for reviewing and improving employees' ethical behavior

Recommendations for Federal Employees:

- Stay up to date with responsibilities and conduct oneself in accordance with the legal authorities, professional codes of conduct, and ethical standards of their organization
- Perform data activities in accordance with the legal, professional, and ethical standards that apply to their areas of work and types of data used
- Understand the policies for handling different types of data within their organization
- Take training to support ethical acquisition, management, and use of data as it aligns to their data role and level within the organization

Legal Authorities

- All legal authorities cited herein support this Tenet

Additional Resources

- American Statistical Association (ASA) Ethical Guidelines for Statistical Practice
- Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
- Office of Government Ethics – 14 General Principles
- Principles and Practices for a Federal Statistical Agency

Figure 2: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.1 - Be Aware of and Uphold Applicable Statutes, Regulations, Professional Practices, and Ethical Standards

4.2 Be Honest and Act with Integrity

All federal leaders and employees, regardless of job title, specific data responsibilities, and role in the organization are expected to exhibit honesty and integrity in their work with data. Data professionals should not perform or condone unethical data behaviors. When sharing data and findings, data professionals should accurately report information and present data limitations, any known biases, and methods of analysis. Data professionals should also recognize that no dataset can fully represent all facets of a person, community, or issue. Data professionals are expected to use data to analyze and solve problems, have humility when presenting data, and be open to feedback and invite discussion with the public. Finally, to embody honesty and integrity, data professionals should accurately represent their abilities when working with data, and not take on data roles for which they are not qualified.

Federal agencies should also support honesty and integrity by clearly defining processes for reporting data ethics concerns and violations, and federal leaders and staff should appropriately implement those processes. It is recommended that each agency develop and communicate policies and procedures to protect those reporting issues.

Recommendations for Federal Leaders:

- Develop a culture of honesty and integrity within their organizations, setting the example of ethical data acquisition, management, and use for their colleagues to follow
- Assign qualified personnel to conduct program data activities
- Implement policies and procedures for the appropriate use of data by agency employees
- Understand and disclose any known limitations, defects, or biases
- Provide mechanisms for employees to anonymously report ethical violations with data
- Use techniques to limit bias during data collection
- Strive for objective analysis

Recommendations for Federal Employees:

- Conduct data activities with honesty and integrity, and in accordance with organizational policies and procedures
- Only conduct data responsibilities for which one is qualified
- Document the data collection and curation process to promote an understanding of the data used in analysis and reporting and to enable reproducibility of results
- Document and communicate the data lineage to promote understanding of where the data came from, how the data was used, and who used it
- Document and communicate any known data limitations, defects, or biases that could not be mitigated during data collection
- Adhere to organizational policies and procedures to report when unethical behaviors are witnessed or suspected
- Use established methods and protocols to limit bias during data collection wherever possible
- Strive for objective analysis

Note: Legal Authorities and Additional Resources included on Page 9

Legal Authorities:

- Executive Office of the President, Office of Science and Technology Policy (OSTP) Scientific Integrity Policy
- OMB M-14-06, Guidance for Providing and Using Administrative Data for Statistical Purposes

Additional Resources:

- American Statistical Association (ASA) Ethical Guidelines for Statistical Practice
- Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
- Department of Homeland Security (DHS), Compliance, Computer Matching Programs
- Department of Education, National Center for Education Statistics (NCES), The Forum Guide to Data Ethics
- Federal Policy for Protection of Human Research Subjects ('Common Rule')
- OMB Questions & Answers When Designing Surveys for Information Collections
- Principles of Artificial Intelligence Ethics for the Intelligence Community
- Artificial Intelligence Ethics Framework for the Intelligence Community

Figure 3: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.2 – Be Honest and Act with Integrity

4.3 Be Accountable and Hold Others Accountable

Accountability requires that anyone collecting, managing, or using data be aware of data stakeholders and be responsible to them as appropriate. This includes upholding data use agreements made with data providers, obtaining informed consent before using data for purposes other than the original intent, and allowing for public access, amendment, and contestability to data and findings, where appropriate. The list of data stakeholders below is not exhaustive, but represents common groups to which anyone working with data might be responsible:

- Individuals and communities who provide data as respondents or serve as research subjects
- Those directly impacted by data use, such as recipients of program services
- Members of the public who rely on data products
- Data consumers, including customers or clients requesting data, who may be a person internal (e.g., agency program manager) or external to the agency (e.g., Congressional staffer)

Recommendations for Federal Leaders:

- Provide data ethics training and skills analysis to any persons working with, interpreting, and communicating data findings
- Assign accountability to specific individuals for ethical considerations through the data lifecycle
- Establish procedures that allow for public access, amendment, and contestability to data and findings, where appropriate
- Maintain data governance policies and practices over time, updating them when necessary
- Implement data sharing and use agreements, such as Memoranda of Understandings (MOU), Inter-Agency Agreements (IAA), and contracts when needed and as appropriate, and with ethical principles reflected in the terms and conditions
- Document processes for data activities and decisions to enable accountability, auditing, and oversight
- Consider providing centralized guidance of data ethics within agencies

Federal Employees should:

- Consider stakeholders while conducting data activities and determine appropriate engagement, keeping their interests in mind and upholding the public trust
- Take data ethics and skills training to improve ability to work with, interpret, and communicate data findings
- Allow for public access, amendment, and contestability to data findings, where appropriate and in accordance with organizational policies
- Uphold data governance policies and data ethics standards practices
- Uphold data use agreements made with data providers
- Document how data are collected, curated, and analyzed for accountability purposes

Note: Legal Authorities and Additional Resources included on Page 11

Legal Authorities

- Foundations of Evidence-Based Policymaking Act of 2018
- Information Quality Act (IQA) of 2001
- Improving Implementation of the IQA, OMB Memorandum M-19-15
- OMB Circular A-130, "Managing Information as a Strategic Resource"
- Paperwork Reduction Act of 1995

Additional Resources

- United States Census Bureau's Data Stewardship Executive Policy Committee
- NIST Privacy Framework
- Utrecht Data School Data Ethics Decision Aid (DEDA)
- United States Geological Survey's Data Sharing Agreements

Figure 4: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.3 - Be Accountable and Hold Others Accountable

4.4 Be Transparent

Individuals and communities benefit when the ethical decision-making process is as transparent as possible to stakeholders. Transparency is grounded in clear communication of all aspects of data activities and appropriate engagement with data stakeholders. This requires engaging stakeholders through easily accessible feedback channels and providing timely updates on the progress and outcomes of data use.

Recommendations for Federal Leaders:

- Develop a culture that supports appropriate transparent reporting of their organization's data activities and products
- Establish standards and provide training for data preparation, documentation, and presentation to promote accuracy and consistency, as well as improved understanding by stakeholders
- Promote clear guidance that ensures data are made available for research equitably and objectively
- Implement standards to clearly document descriptions of analytical methods and models to be shared with appropriate stakeholders
- Establish procedures for making corrections to previous reporting that might contain errors, explaining what was inaccurate and corrected if feasible

Federal Employees should:

- Take training as appropriate for data preparation and presentation
- Follow standard data preparation and presentation methodologies
- Make data available for research in an equitable and objective manner
- Clearly document data activities in ways that can be communicated to stakeholders
- Accurately document metadata, as necessary
- Follow procedures to correct previously reported data that might contain errors, providing clear explanation of what was inaccurate and how it was corrected if feasible

Legal Authorities

- Foundations of Evidence-Based Policymaking Act of 2018
- Information Quality Act (IQA) of 2001
- Improving Implementation of the IQA, OMB Memorandum M-19-15
- OMB Statistical Policy Directive No. 4: Release and Dissemination of Statistical Products Produced by Federal Statistical Agencies
- Paperwork Reduction Act of 1995
- Privacy Act of 1974
- The Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002
- The Plain Writing Act of 2010

Additional Resources

- Commission on Evidence-based Policymaking (CEP) Final Report: The Promise of Evidence-Based Policymaking

Figure 1: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.4 - Be Transparent

4.5 Be Informed of Developments in the Field of Data Science, Including with Data Systems, Techniques, and Technologies

Advanced technologies provide great benefit to the public sector but should be deployed with a commitment to accountability and risk mitigation. While traditional data use and analysis can introduce bias, emerging systems, technologies, and techniques require additional awareness and oversight.

First, emerging systems, technologies, and techniques can be used by those who do not have the requisite training or experience to perform the work. Second, advanced technologies and computational methods can lead to algorithms and automations that make probabilistic recommendations and decisions on behalf of programs (e.g., acceptance of application for benefit program), potentially impacting many individual lives and businesses with minimal human oversight.

Therefore, it is important to remain informed of developments in the field of data science, especially as those methods impact future data collection, management, and use. Examples of advanced technologies, analytics, and computational methods that should be monitored and assessed include:

- Artificial Intelligence (AI)
- Machine Learning (ML)
- Neural Networks (NNs)
- Robotic Process Automation (RPA)
- Internet of Things (IoT)
- Blockchain

In addition, since new data innovations (e.g., systems, solutions, computational methods) emerge every day, it is important for federal leaders and employees working with data to keep abreast of those innovations and learn how to ethically use those methods.

For Those Using AI, ML, Advanced Analytics, and Computing Technologies:

The Framework does not offer specific guidance on advanced technologies and computational methods. The Framework is generally applicable to those methods, however, and references other federal initiatives that more thoroughly address the ethical considerations associated with the deployment of such methods.

In support of each Tenet, the Framework encourages those who use AI, ML, advanced analytics, and computing technologies to:

- Design, develop, and use those methods while respecting the law and acting with integrity
- Provide appropriate transparency around methods, applications, and uses
- Understand and disclose any known limitations, defects, or biases
- Develop and employ mechanisms to identify responsibilities and provide accountability
- Institute rigorous protocols to evaluate outputs for bias and implement mitigation strategies as needed
- Prioritize human-centered development and use
- Leverage mechanisms to communicate the design, development, and use process to non-technical stakeholders
- Engage other scientific and technology communities to leverage best practices
- Seek out specific ethical guidance around specific emerging techniques

Figure 6: Recommendations for those using AI, ML, Advanced Analytics, and Computational Technologies

Recommendations for Federal Leaders:

- Develop a diverse workforce to support the policies, oversight, and governance structure for any large-scale system that learns from data to limit bias and to best consider societal and cultural consequences
- Provide or support training in data science and any new systems, technologies, or techniques if required for job function
- Employ advanced methods in ways that fully comply with applicable legal authorities, policies, and procedures that protect privacy, civil rights, and civil liberties
- Monitor advanced methods and how they might impact their organization's data activities
- Hold employees responsible for staying abreast of advanced methods and how they might be used for data activities in their areas of work
- Establish protocols explicitly designed to identify and mitigate bias, as well as assign accountability, when designing, developing, and deploying advanced methods
- Ensure human involvement in development and use of advanced methods

Recommendations for Federal Employees:

- Take training or develop required knowledge in data science and any systems, techniques, and technologies before applying in the field
- Deploy advanced methods in ways that fully comply with applicable legal authorities and organization policies and procedures
- Keep abreast of advanced methods and how they might impact data activities in their areas of work
- Promote accountability and properly mitigate risks when designing, developing, and deploying advanced methods
- Involve human judgement when advanced methods might interfere with privacy, civil rights, or civil liberties

Legal Authorities

- Executive Order on Maintaining American Leadership in Artificial Intelligence (AI)

Additional Resources

- Association for Computing and Machinery (ACM) Code of Ethics and Professional Conduct
- Australian Government – Artificial Intelligence: Australia's Ethics Framework
- Deon Data Science Ethics Checklist
- General Services Administration (GSA) AI Center of Excellence (CoE)
- Government of Canada – Algorithmic Impact Assessment
- McKinsey & Company – Confronting the Risks of Artificial Intelligence
- Principles of Artificial Intelligence Ethics for The Intelligence Community
- Artificial Intelligence Ethics Framework for the Intelligence Community

Figure 7: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.5 - Be Informed of Developments in the Field of Data Science, Including with Data Systems, Techniques, and Technologies

4.6 Be Respectful of Privacy and Confidentiality

Privacy (i.e., the state of being free from unwarranted intrusion into the private life of individuals) and confidentiality (i.e., free from inappropriate access and use) should always be protected in a manner that respects human dignity, rights, and freedom, while also being in accordance with the applicable legal authorities. An essential objective of privacy and confidentiality assurance is to minimize potential negative consequences.

Confidential information obtained by agencies should always be protected by the appropriate access, use, and disclosure limitations. Data activities that involve individual privacy should fundamentally align with the Fair Information Practice Principles (FIPPs), which are:

- Transparency
- Individual Participation
- Purpose specification
- Data minimization
- Use Limitation
- Data quality and integrity
- Security
- Accountability and auditing

Recommendations for Federal Leaders:

- Provide training to employees on appropriate handling of sensitive data
- Monitor technological advances that increase the risk of identification of individuals or entities represented in public datasets and regularly update disclosure protection protocols to mitigate these risks to the greatest extent possible
- Make clear the tradeoff between confidentiality and granularity of data in public data releases and, to the extent possible, provide tools to enable users to evaluate the impact data protection measures have on results obtained from the data
- Establish protocols for notifying data providers if there is a breach that potentially impacts their privacy and provide mechanisms for accommodating victims of those breaches
- Support and implement mechanisms that limit privacy and confidentiality risks, such as disclosure limitations and controlled access to data
- Comply with applicable legal authorities that govern the protection and use of sensitive data, establishing policies and procedures to prevent re-identification of sensitive data made public, maintain the minimum amount of sensitive data necessary, and adhere to data sharing and use agreements

Federal Employees should:

- Appreciate the rights and responsibilities related to the collection, management, and use of sensitive data
- Take training to ensure the appropriate handling of sensitive data during everyday activities
- Take appropriate measures to comply with organizational policies and procedures to prevent re-identification of sensitive data made public, maintain the minimum amount of sensitive data necessary, and adhere to data sharing and use agreements throughout the data lifecycle
- Only use sensitive data for authorized purposes and without violating protection assurances
- Only access data with appropriate authorization or approval

Note: Legal Authorities and Additional Resources included on Page 16

Legal Authorities

- Defend Trade Secrets Act of 2016
- E-Government Act of 2002
- Department of Homeland Security (DHS), the Fair Information Practice Principles: Framework for Privacy at the DHS
- Federal Information Security Modernization Act (FISMA) of 2014
- Foundations of Evidence-Based Policymaking Act of 2018
- Freedom of Information Act (FOIA)
- Information Quality Act (IQA) of 2001
- Improving Implementation of the IQA, OMB Memorandum M-19-15
- Federal Policy for Protection of Human Research Subjects ('Common Rule')
- M-03-22, OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002
- OMB Circular A-130, "Managing Information as a Strategic Resource"
- Presidential and Federal Records Act Amendments of 2014
- Privacy Act of 1974
- The Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002
- Title 13, United States Code (U.S. Census Bureau)

Additional Resources

- NIST Privacy Framework

Figure 8: Recommendations and Resources for Federal Leaders & Employees for Tenet 4.6 - Be Respectful of Privacy and Confidentiality

4.7 Be Respectful of the Public, Individuals, and Communities

Data activities have the overarching goal of benefiting the public good. All other Tenets support this notion, and responsible federal leaders and employees approach data activities with promoting the "public good" in mind. Responsible use of data begins with a careful consideration of its potential impacts. As a result, data initiatives should have an identified and clear benefit to society that guide their execution.

In situations where data on individuals or communities are used to deliver services to the public, it is a best practice to engage with impacted stakeholders to better understand those represented by the data.

Recommendations for Federal Leaders:

- Promote the protection of privacy, civil rights, and civil liberties in their organization's data use
- Carefully consider the impacts their organization's data activities might have on the public, individuals, and communities, taking measures to minimize any negative consequences
- Where negative consequences are unavoidable, establish procedures to mitigate harm
- Assess stakeholders and implement procedures to appropriately engage those impacted by the organization's data activities

Federal Employees should:

- Understand that their data activities might impact the public, individuals, and communities
- Strive to promote the public good through their data activities
- Holistically consider the impacts their data activities might have on the public, individuals, and communities, and take measures to minimize any negative consequences
- Where negative consequences are unavoidable, take measures to mitigate harm
- Where data on individuals or communities are used to deliver services to the public, engage those represented by the data to better understand and promote their interests

Legal Authorities

- Federal Information Security Modernization Act (FISMA) of 2014
- Foundations of Evidence-Based Policymaking Act of 2018
- Information Quality Act (IQA) of 2001
- Improving Implementation of the IQA, OMB Memorandum M-19-15
- Privacy Act of 1974

Additional Resources

- American Statistical Association (ASA) Ethical Guidelines for Statistical Practice
- Association for Computing and Machinery (ACM) Code of Ethics and Professional Conduct

Figure 9: Recommendations and Resources for Federal Leaders & Employees for 4.7 - Be Respectful of the Public, Individuals, and Communities

5 Data Ethics Tenets in Action

5.1 Benefits of Data Ethics

The Data Ethics Framework helps guide the data activities of agencies, providing the foundation for the ethical acquisition, management, and use of data for any federal purpose. An explanation of the Framework's benefits includes:

- **Consistency.** All federal employees reference an agreed-on, principles-based set of Tenets to help navigate the ethical considerations of data use. Agency personnel from different domains and fulfilling different roles in the government apply the same foundational ethical considerations.
- **Better, Data-Driven Decisions.** Support data methods and processes that uncover data limitations, gaps, and biases. Make justifiable decisions and communicate known data limitations to promote transparency. Encourage data users to actively consider data limitations and gaps at all stages of analysis, especially when making decisions with the data.
- **Risk Mitigation.** From the beginning of the data initiative, identify and assess potential impacts on individuals, organizations, and society as a whole. Leverage the right technologies to appropriately store data, secure data, maintain complete metadata and documentation, understand data lineage, improve data quality, manage data, and appropriately dispose of data.
- **Increased Transparency.** Agencies ensure the documentation and communication of trustworthy data processes, increasing the transparency around federal data collection, testing, use, and dissemination.
- **Consideration of Wider Perspectives.** Promote collaboration across internal and external stakeholder groups to better understand data subjects and impacts of data use.
- **Improved Public Trust.** Engender public trust through comprehensive stakeholder engagement, ensuring accountability across the data lifecycle, and reinforcing protocols to protect privacy, civil liberties, and civil rights, as well as confidentiality.

The way data is used continues to touch almost every aspect of daily life. Although the ethical challenges that come with data use are many, integrating the Framework's guidance into everyday agency activities will help mature data ethics considerations – and with it, its benefits – across the Federal Government.

5.2 Use Cases

The Use Cases herein provide examples of ethical considerations with data encountered while doing federal work. They demonstrate the application of the Framework across the data lifecycle and how to use the Framework's recommendations as a reference for ethical decision making.

The questions included are examples of the important considerations for federal leaders and employees before, during, and after data initiatives. They are not meant to be all inclusive. In addition, it is recognized that ethical use of data is supported by the ethical behavior of organizations. The Framework's Use Cases aim to complement existing cultural and training initiatives that reinforce ethical behavior across the government.

5.2.1 Use Case: Artificial Intelligence & Bias

Artificial Intelligence & Bias	
Organization Type:	Government Benefit Agency
Primary Program Objective:	Improve administration of government benefits
Secondary Project/Program Objective:	Leverage administrative data and advanced technologies (i.e., AI) to improve performance
Scenario:	An agency that oversees administration of benefits collects large amounts of applicant data on a daily basis. To streamline the application process, the agency engaged an outside party to create an automated tool that makes decisions on applicant eligibility for the benefits program. The tool relies on models that gather data from different parts of the organization, including applicant employment and financial records, and analyzes the chances the applicant will be successful. In operation for over two years, the tool helps eliminate numerous manual processes, identify potential fraud, and better deploy limited resources. During this time, the tool has made thousands of decisions on applicant eligibility for benefits, impacting countless lives. Karen, who works for the agency's outreach department, has received an increasing volume of complaints from applicants in recent months. Applicants have consistently stated they have been inappropriately screened out of the application process. Karen brings the issue to her management, who demand an internal review. The internal review finds that data sharing agreements with the outside party who created the automated tool required for the underlying data and code to be destroyed upon the tool's deployment on the agency's customer-facing website. As a result, the agency is unable to reproduce, evaluate, or scrutinize the tool and its supporting decision models.
Use Case Questions:	• Was the agency justified in using applicant data to improve their own processes via automated decision models? • Could altered data sharing agreements helped to enable reproducibility of the impactful decision models? • What ethical considerations should have arisen during the design, development, and deployment of the automated tool?

Data Lifecycle Questions:

Creation or Collection (Acquisition)

- Are there any limitations to the data the agency is allowed to collect for program administration purposes?
- Should members of the public know how data collected during the application process is ultimately used?
- Should all data being collected serve a distinct purpose?
- How could bias in the collection process be mitigated?

Processing

- Are there issues with leveraging data from other parts of the organization (i.e., employment and financial records) for model development?
- Should there be oversight measures in place to ensure the correct records are linked across the organization?
- Should processing activities and effects on data quality documented?

Dissemination

- What should have been considered when arranging the data sharing agreement with the outside party?

Use:

- Is the appropriate level of human judgement involved in decisions produced by the automated tool?
- Should usage of data produced by the automated tool be monitored?
- Should the probabilistic recommendations from the automated tool be reviewed and/or validated?
- Should applicants be aware that their personal data are being used to drive automated decisions on benefits and allowed to contest those decisions?

Storage and Disposition:

- Should the underlying data and code supporting the automated tool have been stored for accountability purposes?
- Should the agency have considered documenting analytical methods and results in this scenario?
- Should the agency have procedures in place to dispose of the data after a certain timeframe?

5.2.2 Dissemination & Impacts

Dissemination & Impacts	
Organization Type:	Government Inspection Agency
Primary Program Objective:	Enforce the law and minimum standards of the Animal Welfare Act
Secondary Project/Program Objective:	Publish reports to increase the public's understanding of adherence to the act's standards
Scenario:	The Animal Welfare Act (AWA) regulates the treatment and care for certain animals bred for commercial sale, used in research, transported commercially, or exhibited to the public. The government inspection agency enforces the law, setting minimum standards as the baseline by which they assess a facility's care for animals. Each year, the agency collects, manages, and analyzes AWA records from animal facilities (e.g., animal shelters, retailers) to support its mission and perform reporting. In addition, the agency posts information and reports from investigations and inspections to a public website. The reports give the public a picture of the compliance of all entities licensed or registered under the AWA. Due to a pending lawsuit and privacy concerns, the government inspection agency removed certain reports and AWA records from its website that identified animal abuse complaints (i.e., complaints not convictions). Users, activists, and the public were upset, as posting the AWA records not only seems like the right thing to do, but the agency has taken measures to redact some information previously posted. The agency communications department is in a tight spot – there is a demand for public transparency, but legal matters and a transition in agency leadership leave the team with conflicting views on how to proceed.
Use Case Questions:	• Is the agency justified in removing certain reports and AWA records from its website, instead of trying to redact certain private information? • Should the agency have disclosure limitation methodologies in place to prevent this type of situation? • How could the agency weigh consumers right to know potential animal abuse violations and also protect privacy of those not convicted? • Should the agency have policies and procedures in place to provide guidance in this situation, despite the transition in agency leadership?

Data Lifecycle Questions:

Creation or Collection (Acquisition)

- How will data quality be checked and be documented?
- Who needs to be involved in the data collection?
- Have the entities (i.e., animal shelters, retailers) been informed of the data to be collected and how it will be used?

Processing

- Once acquired, what refinement does the data require before release? For example, what private or confidential information should be removed?
- What considerations need to be weighed in protecting privacy while analyzing the data?
- What additional information around data collection, maintenance, and use needs to be shared to promote public transparency?

Dissemination

- What data need to be provided to the public?
- How can data be shared with external stakeholders?
- Are there open feedback channels for stakeholders to share insights or to report incorrect information?
- Are the data presented in a consumable way?

Use

- Could the data be used by data consumers in a nefarious or harmful way?
- What types of decisions or uses could known data consumers make with the data?
- Could the data be joined with other data to identify personal identities or confidential information?

Storage and Disposition

- How long should this data be maintained?
- Does the data storage mechanism document any changes in data collection or curation that could impact the long-view of the data?