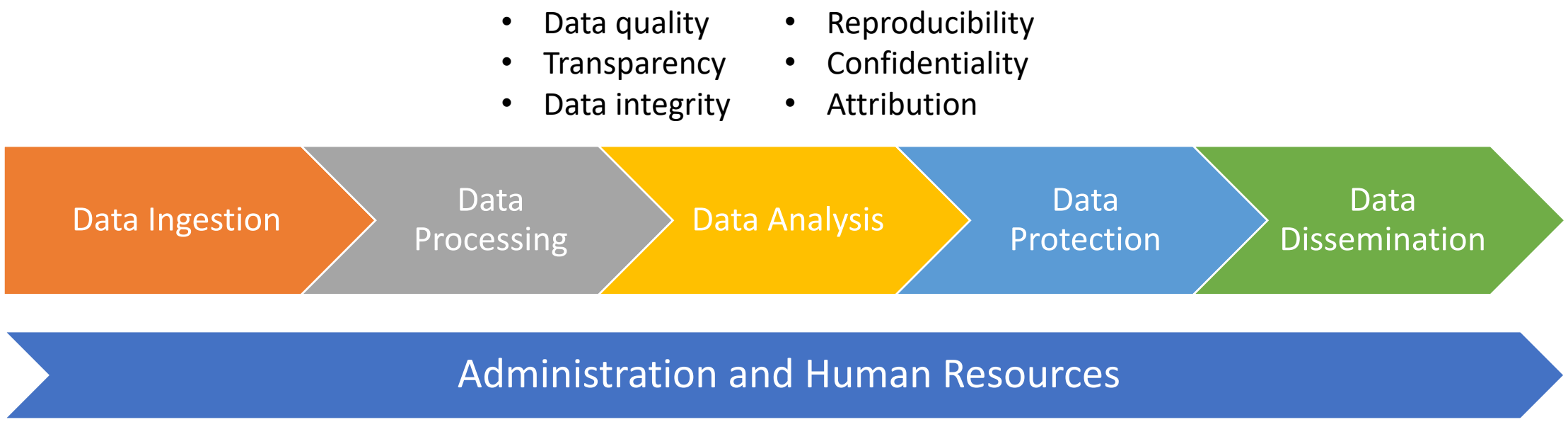
Ethical Considerations of AI for the Federal Statistical System

Michael Hawes
Senior Statistician for Scientific Communication
U.S. Census Bureau
michael.b.hawes@census.gov

AI Day
May 2, 2024
National Academies of Science, Engineering, and Medicine
Ethical Implications of AI for Statistical Agencies Breakout Session

Ethical Challenges of AI for Federal Statistical Agencies



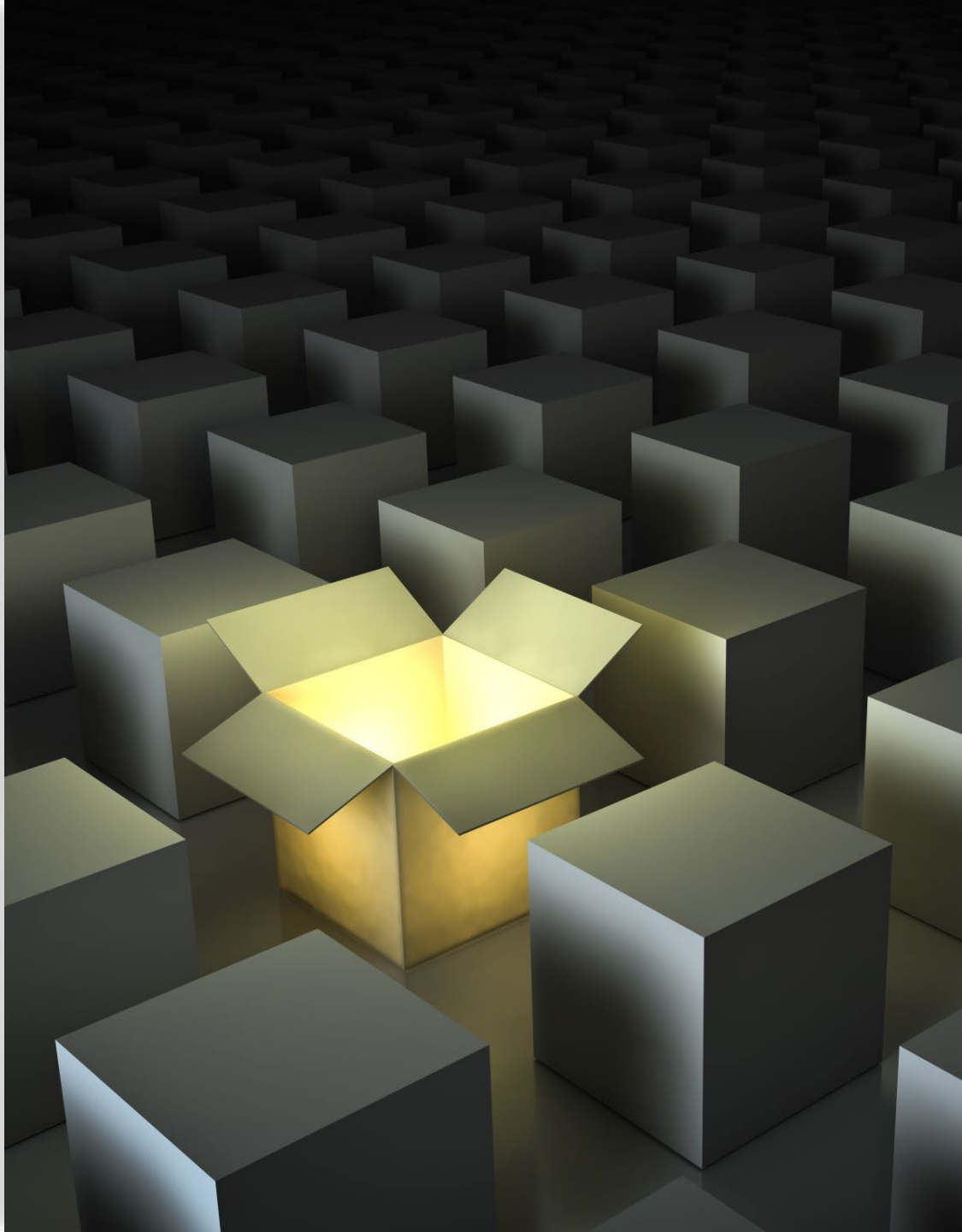
Plus Ça Change…Plus C'est la Même Chose

- Many of the ethical challenges of AI are not truly new…just more pronounced.

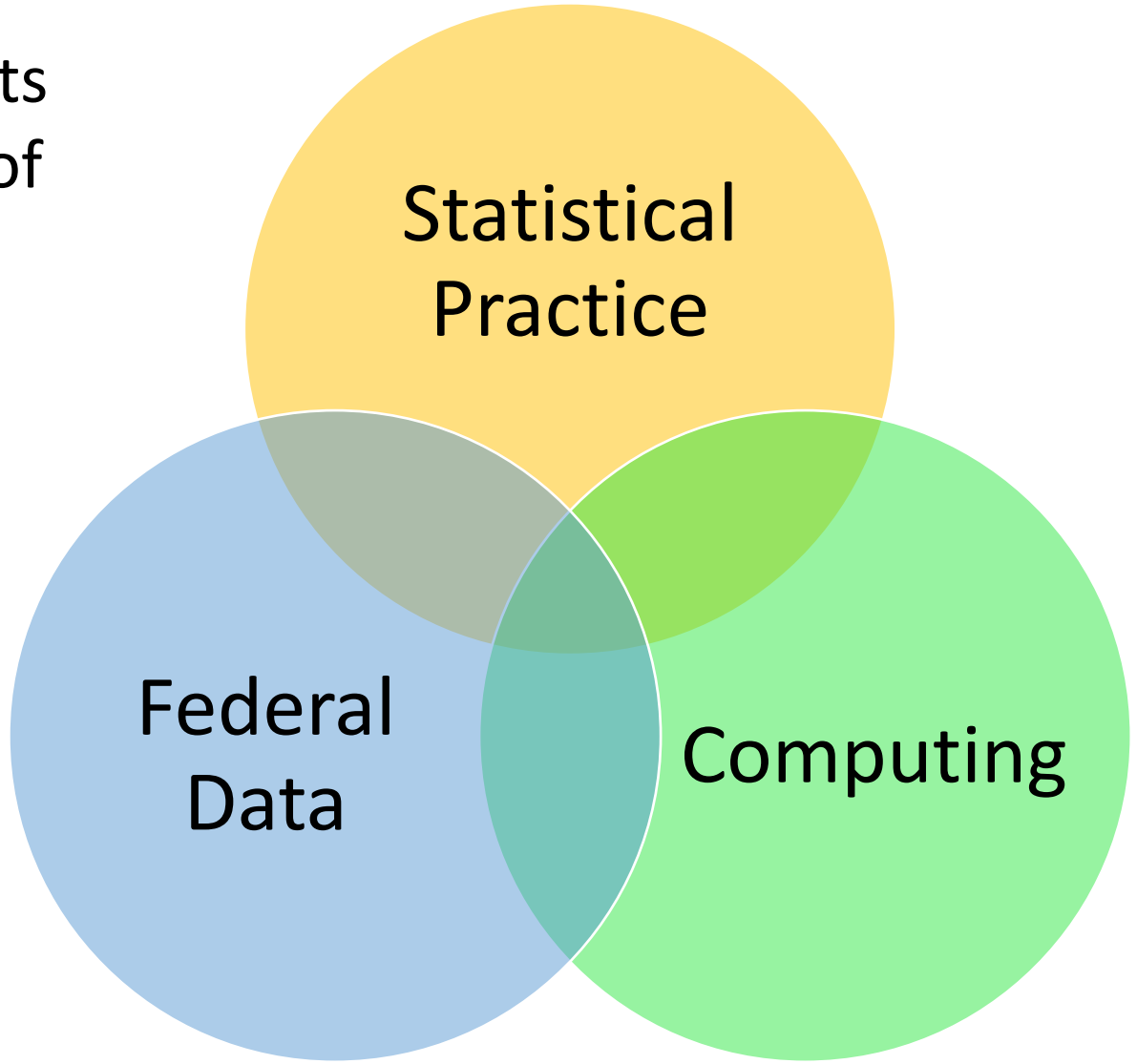
  (e.g., data quality for non-traditional data sources; attribution and integrity of statistics served up by third parties)

- Some, however, do mark new and potentially vexing ethical challenges.

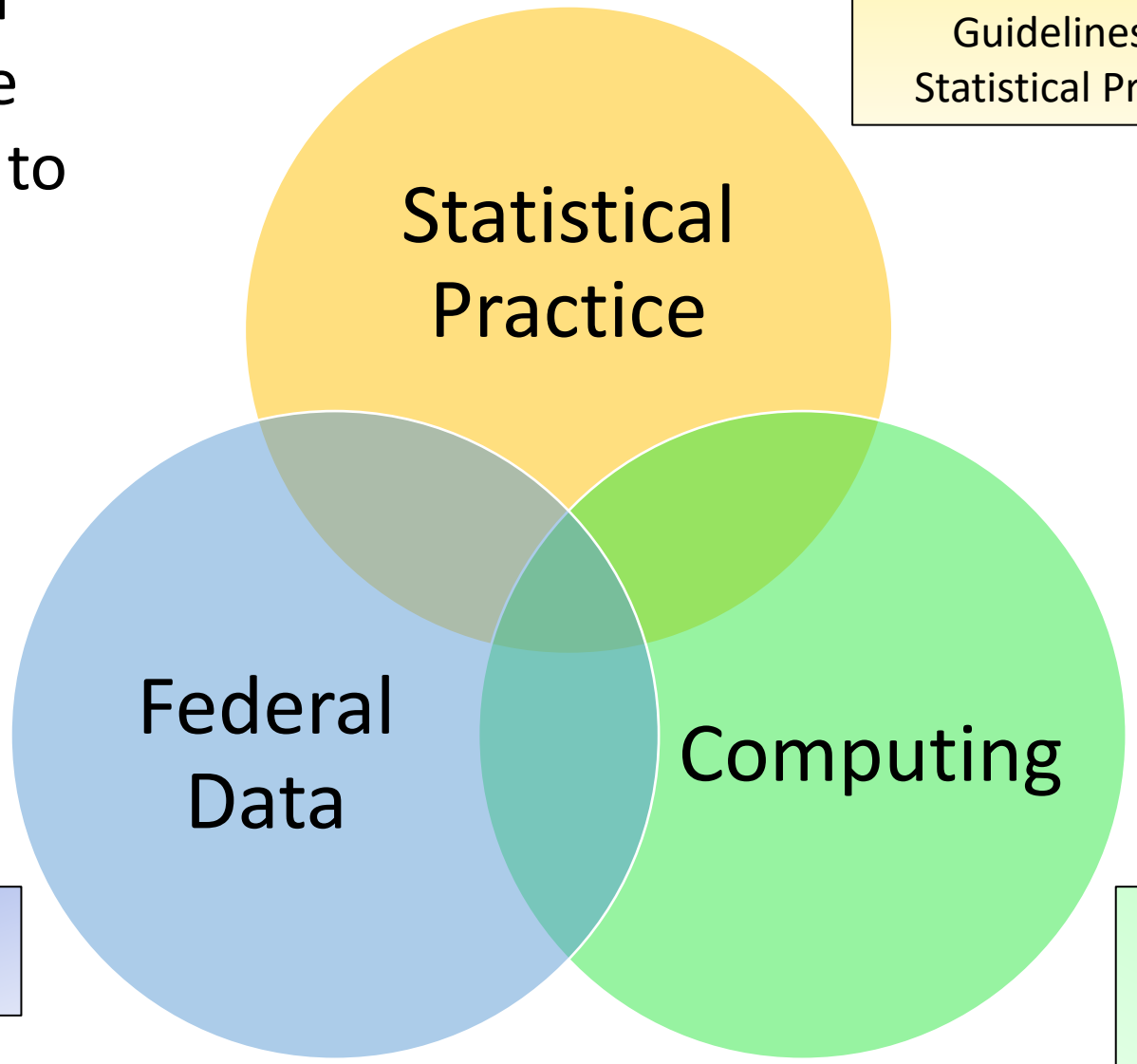
  (e.g., transparency and contestability of "black box" AI applications)



AI in the statistical agency context exists at the intersection of three domains.



Fortunately, each of those domains have ethical frameworks to help us navigate these challenges.

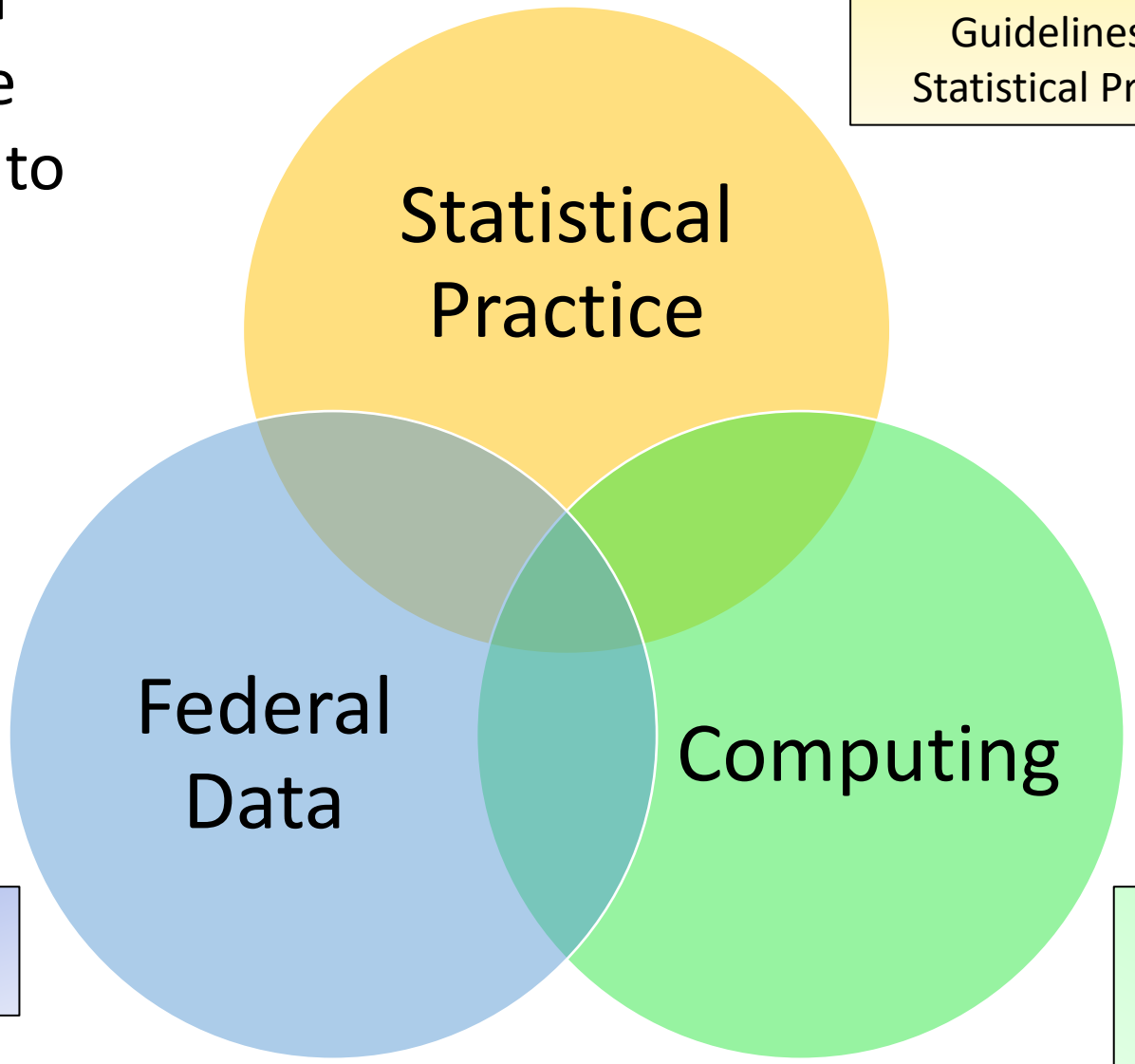


American Statistical Association Ethical Guidelines for Statistical Practice

Federal Data Ethics Framework

Association for Computing Machinery Code of Ethics and Professional Practice

Fortunately, each of those domains have ethical frameworks to help us navigate these challenges.



American Statistical Association Ethical Guidelines for Statistical Practice

Federal Data Ethics Framework

Belmont Report (1978)

Association for Computing Machinery Code of Ethics and Professional Practice

Belmont Report (1978)

Ethical Principles and Guidelines for the Protection of Human Subjects of Research

Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research

Three Guiding Principles:

- **Respect for Persons** *(autonomy, respect, consent, integrity)*
- **Beneficence** *("do no harm", maximize potential benefits and minimize potential harms)*
- **Justice** *(equity)*

Federal Data Strategy

Data Ethics Framework

Background

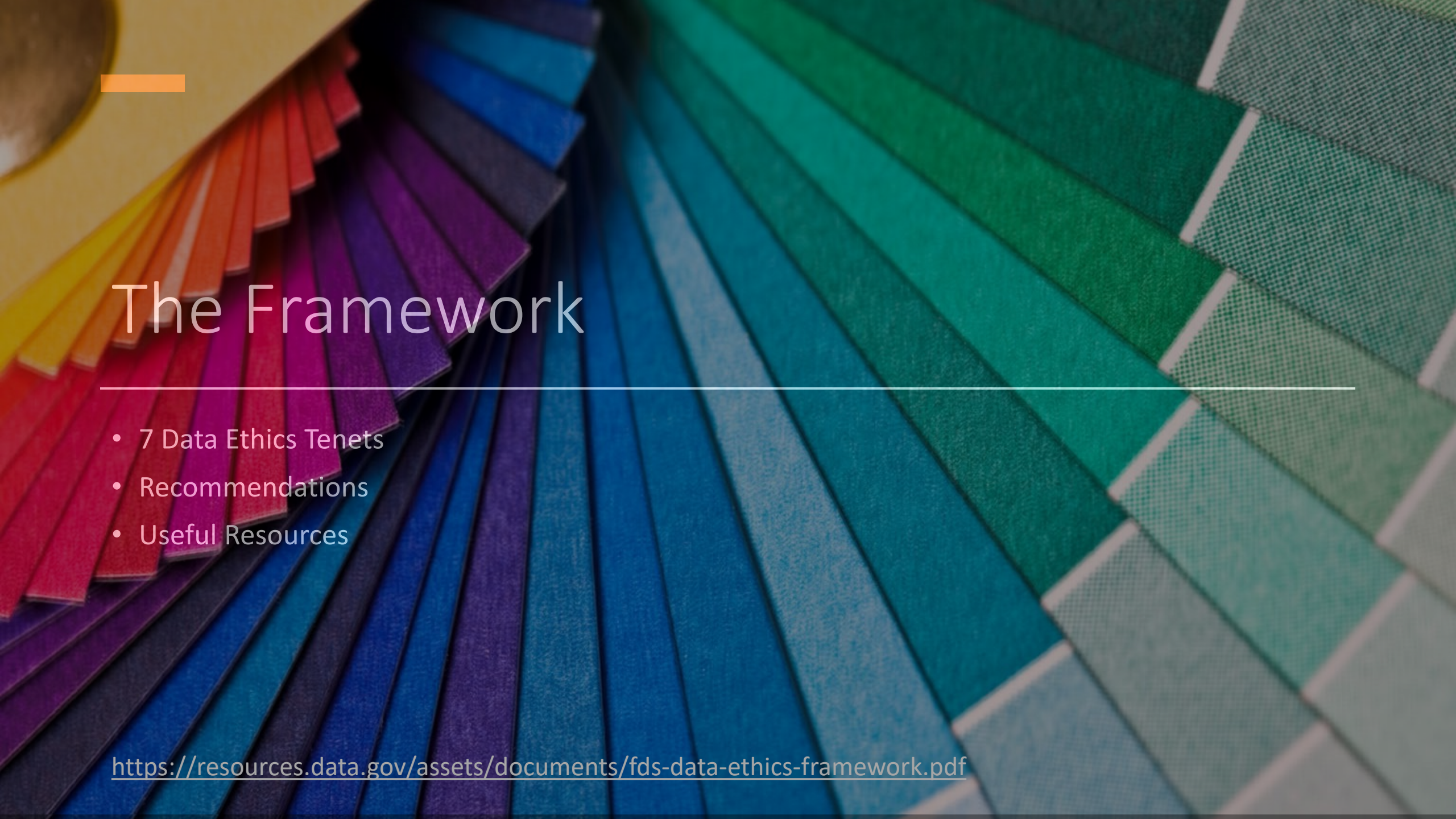
Developed as part of the 2019 Federal Data Strategy, with the support of the Chief Data Officers' Council, the Federal Privacy Council and the Interagency Council on Statistical Policy

Data Ethics Framework Development Team

- **Trey Bradley**, Strategic Data Initiatives Program Manager, Office of Shared Solutions & Performance Improvement, U.S. General Services Administration (Development Lead)
- **Ken Ambrose**, Senior Advisor for the CDO Council, Office of Government-wide Policy, U.S. General Services Administration
- **Maya Bernstein**, Senior Advisor for Privacy Policy, Office of the Secretary for Planning & Evaluation, U.S. Department of Health & Human Services
- **Ivan DeLoatch**, Executive Director, Federal Geographic Data Committee, U.S. Geological Survey, U.S. Department of the Interior
- **Dave Dreisigmeyer**, Interim Deputy Chief Data Officer, Office of the Under Secretary for Economic Affairs, U.S. Department of Commerce
- **Jeffrey Gonzales**, Research Mathematical Statistician, Economic Research Service, U.S. Department of Agriculture
- **Chris Grubb**, Chief Data Scientist, Center for Analytics, U.S. Department of State
- **Lisa Haralampus**, Director of Records Management Policy and Outreach, Office of the Chief Records Officer, U.S. National Archives and Records Administration
- **Michael Hawes**, Senior Advisor for Data Access and Privacy, U.S. Census Bureau, U.S. Department of Commerce
- **Barry Johnson**, Acting Chief, Research and Analytics Office, Internal Revenue Service, U.S. Department of the Treasury
- **Brandon Kopp**, Research Psychologist, Bureau of Labor Statistics, U.S. Department of Labor
- **John Krebs**, Chief Privacy Officer, Federal Trade Commission
- **Justin Marsico**, Chief Data Officer, Deputy Assistant Commissioner, Bureau of the Fiscal Service, U.S. Department of the Treasury
- **Daniel Morgan**, Chief Data Officer, U.S. Department of Transportation
- **Katerine Osatuke**, Research Director, Veterans Health Administration National Center for Organization Development, U.S. Department of Veterans Affairs
- **Eileen Vidrine**, Chief Data Officer, U.S. Air Force, U.S. Department of Defense

Data Ethics Defined

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



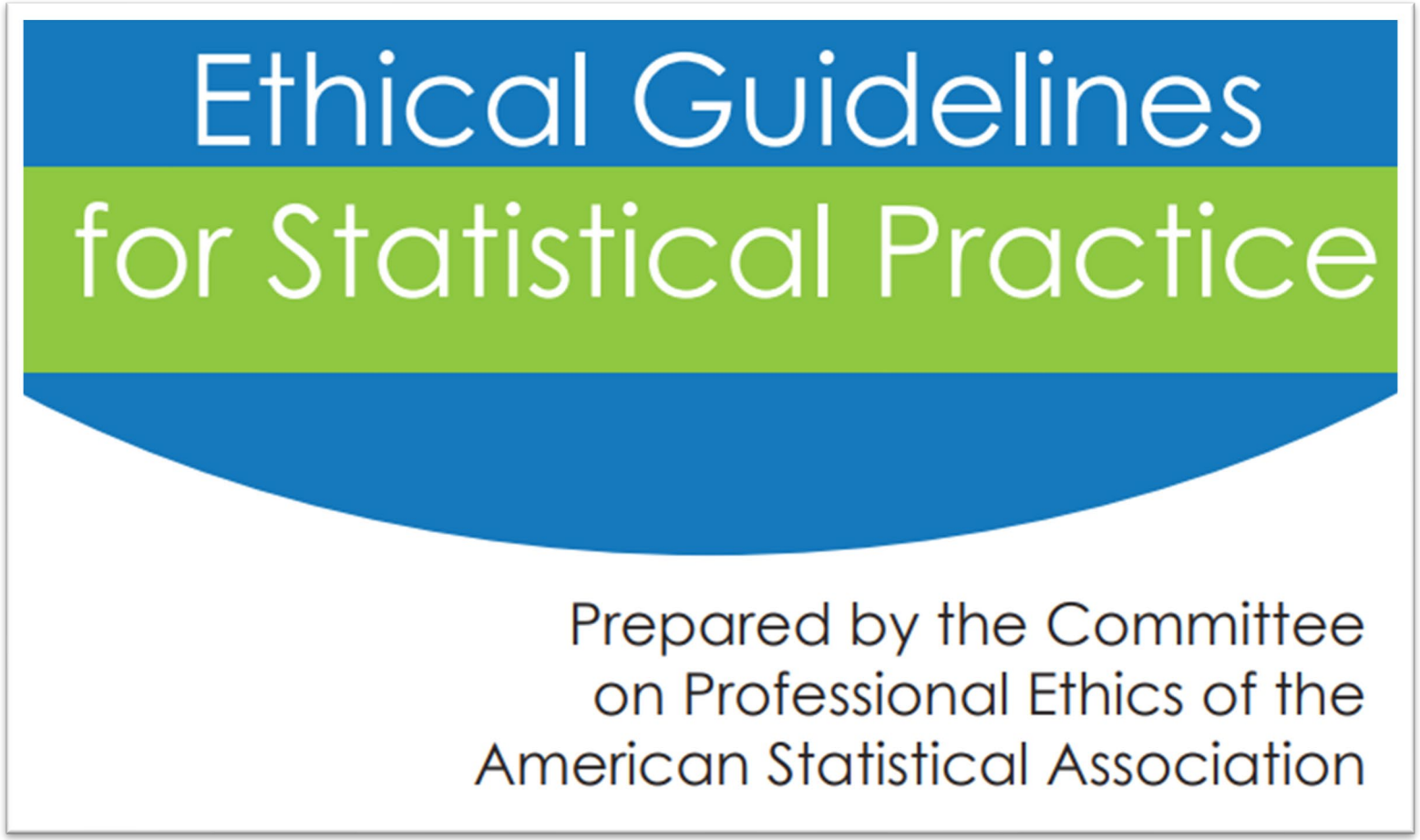
The Framework

- 7 Data Ethics Tenets
- Recommendations
- Useful Resources

https://resources.data.gov/assets/documents/fds-data-ethics-framework.pdf

Uphold Applicable Statutes, Regulations, Professional Practices, and Ethical Standards

- Organizational leaders are encouraged to maintain up-to-date, comprehensive ethical standards regarding data use.
- Staff are responsible for learning and applying agency guidance.
- If a person works in an area with recognized professional ethical codes of conduct (such as computer science or software engineering) they should be aware of those standards and strive to uphold them.



https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice

(2022 Revision)

Respect the Public, Individuals, and Communities

- Responsible federal leaders and data users should approach data activities with promoting the “public good” in mind.
- Responsible use of data begins with careful consideration of its potential and differential impacts.
- Data initiatives should include considerations for unique community and local contexts, such as for federally recognized Tribes, Alaskan Natives, and local governments, and have an identified and clear benefit to society.

Respect Privacy and Confidentiality

- Privacy and confidentiality should always be protected in a manner that respects the dignity, rights, and freedom of data subjects.
- In this context, privacy is the state of being free from unwarranted intrusion into the private life of individuals; and
- Confidentiality is the state of one's information being free from inappropriate access and use.
- An essential objective of privacy and confidentiality protection is to minimize potential negative consequences, such as the mosaic effect

Act with Honesty, Integrity, and Humility

- Federal leaders and data users should not perform or condone unethical data behaviors.
- When sharing data and findings, personnel should accurately report information and present data limitations, known biases, and methods of analysis that apply.
- They should also take care not to overgeneralize based on available data and recognize that no dataset can fully represent all facets of a person, community, or issue.
- Federal leaders and data users are expected to exhibit humility when presenting data, be open to feedback, and invite discussion with the public.
- In addition, federal data users should accurately and honestly represent their abilities when working with data.
- 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

Hold Oneself and Others Accountable

- Anyone acquiring, managing, or using data should be aware of stakeholders and responsible to them, as appropriate.
- Remaining accountable includes the responsible handling of classified and controlled information, upholding data use agreements made with data providers, minimizing data collection, informing individuals and organizations of the potential uses of their data, and allowing for public access, amendment, and contestability to data and findings when consistent with privacy and national security concerns.

Promote Transparency

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

Stay Informed of Developments in the Fields of Data Management and Data Science

- Advanced technologies provide great benefit to the public sector, but should be deployed with a commitment to accountability and risk mitigation.
- Emerging systems, technologies, and techniques require additional awareness and oversight as they can present new—and sometimes hard to detect—opportunities for bias

Questions?