Federal Data Strategy
Data Ethics Framework

Michael Hawes
Senior Statistician for Scientific Communication

FCSM Research and Policy Conference
Session D4: How Can U.S. Professional Guidelines Inform Federal Statistical Policy?
October 25, 2023

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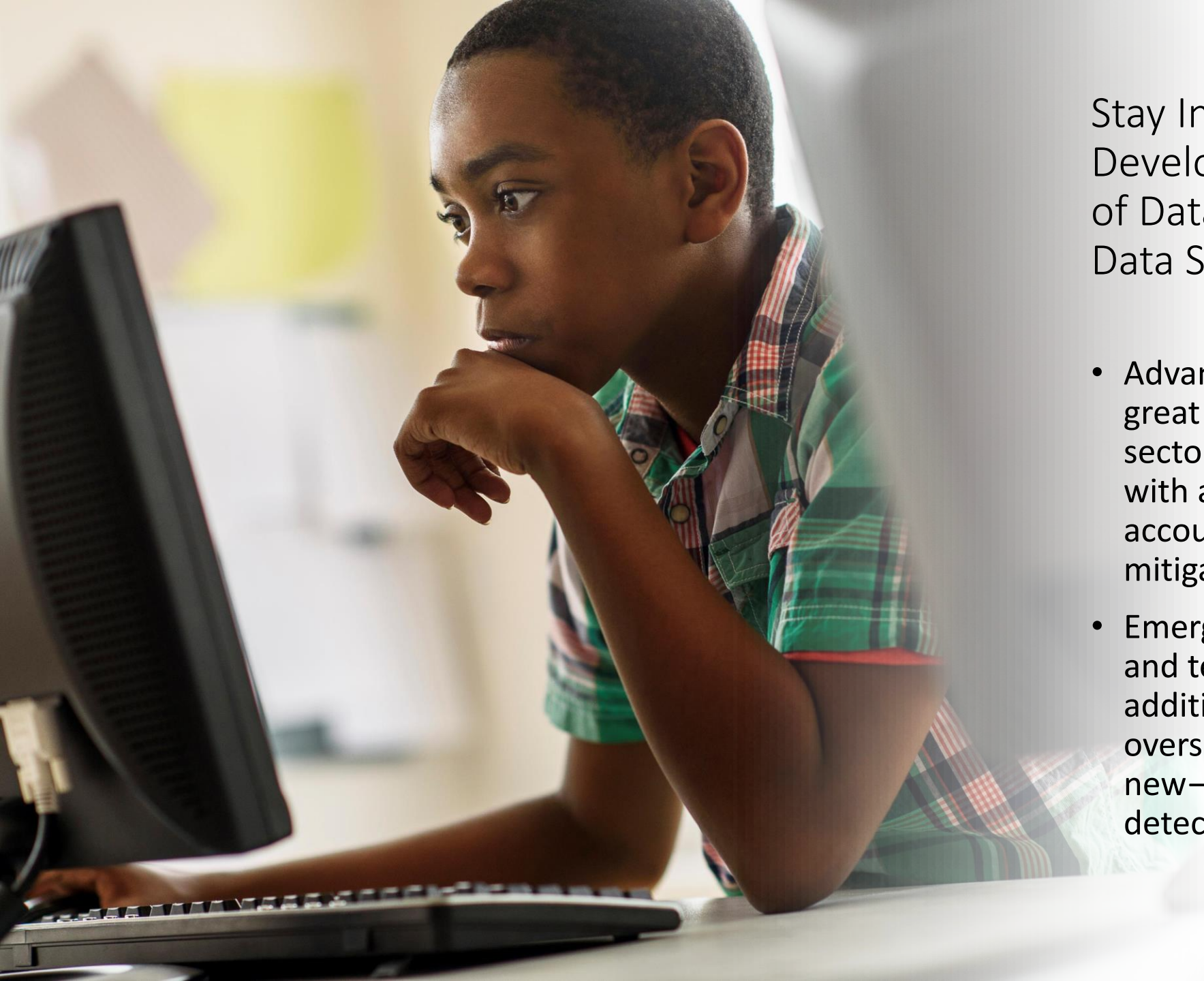
Stay Informed of Developments in the Fields of Data Management and Data Science

- 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
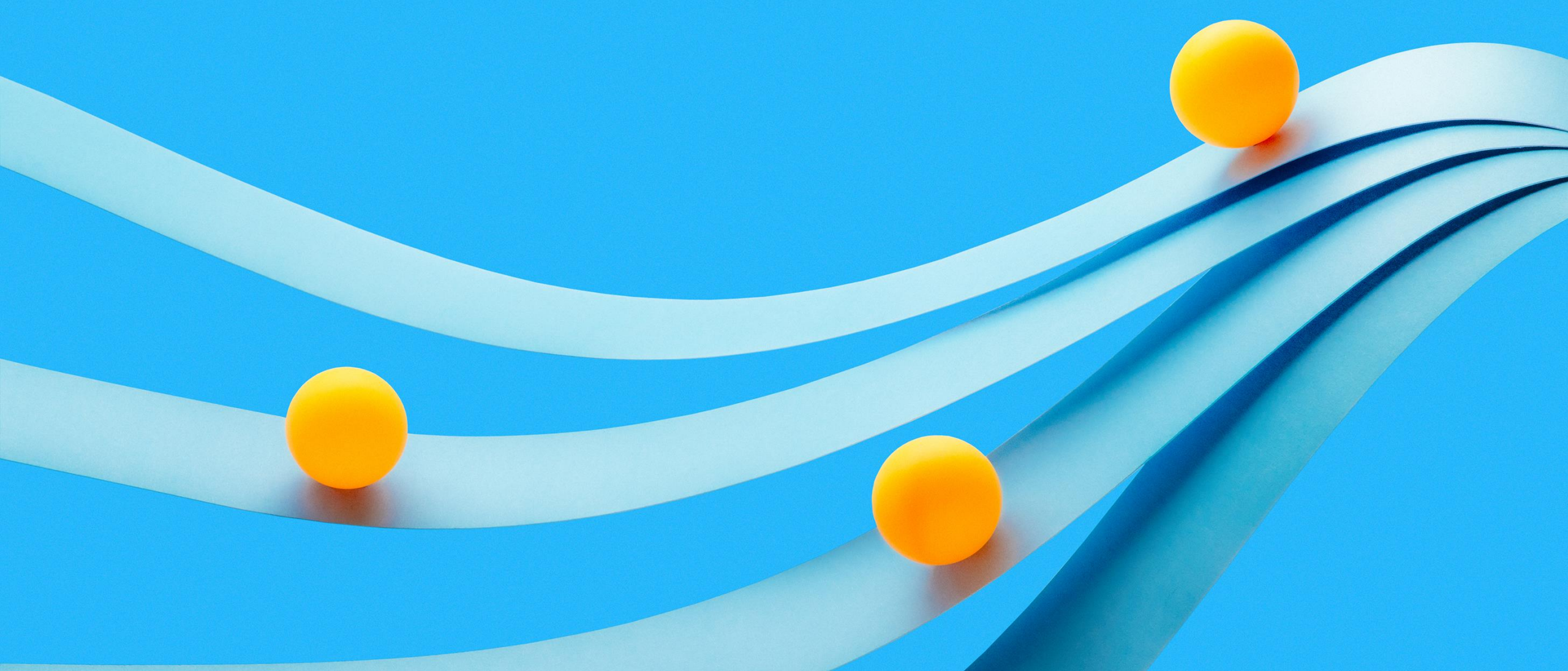
Alignment across frameworks

Results: Example of full table (detail – in Appendix): ASA x Data Ethics Tenets (no red)

TABLE 3: Correspondence between ASA Ethical Guidelines (2022) with US Data Ethics Tenets (2020)

ASA Ethical Guidelines (2022):) / Data Ethics Tenets (2020):	A Professional Integrity and Accountability	B Integrity of Data and Methods	C Stakeholders	D Research Subjects/Data Subjects and Those Affected by Statistical Practices	E Interdisciplinary Team Members	F Other Practitioners/ Profession	G Leader/Supervisor/ Mentor and APPENDIX	H Allegations of Potential Misconduct
1. Uphold applicable statutes, regulations, professional practices, and ethical standards.	A11	(B4)	C2 (C8)	D1, D9, D11	E; E1, E2	F; F1	(G; G1); G2, G5 APPENDIX 1, 2, 4, 7, 10, 11, 12	H; H1, H2, H3, H4
2. Respect the public, individuals, and communities	A; A2, A3, A4, A5, A7	B1, B3, B4, B5, B6	C; C1, C2, C8	D2, D5, D6, D7, D10, D11	E4	F1, F2, F3, F5	G2, G5 APPENDIX 1, 2, 6, 8, 9	H
3. Respect privacy and confidentiality.	(A3)	B4	C7	D4, D5, D7, D9, D10				(H4, H5, H6 - for practitioners- not for data subjects)
4. Act with honesty, integrity, and humility.	A; A1	B1, B2, B3	C1, C2,C3, C4, C8	D10	E2, E4	F3	G; G1, G2, G5 APPENDIX 1, 2, 4, 5, 8, 9, 10, 11	H2
5. Hold oneself and others accountable.	A9, A12		C1, C2	D; D1, D4, D10	E3, E4	F2, F3, F4, F5	G; G1, G2, G5 APPENDIX 1, 2, 4, 12	H2
6. Promote transparency.	A2, A7	B2	C1, C2, C4, C5, C6		E3	F4, F5	G5 APPENDIX 1, 2, 4, 8, 9, 10, 12	(H2)
7. Stay informed of developments in the fields of data management and data science.		B6	(C6)			(F4)	APPENDIX 3 (12)	

Source: Park & Tractenberg (2023)

Results: Summarized table (in body of paper): ASA x Data Ethics Tenets

ASA Ethical Guidelines (2022): Data Ethics Tenets (2020):	A Accountability	B Integrity	C Stakeholder	D Data Providers	E Other Disciplines	F Other Statisticians	G Leadership	H Misconduct
1 Uphold laws and regulations	✓	✓	✓	✓	✓	✓	✓	✓
2 Respect public	✓	✓	✓	✓	✓	✓	✓	✓
3 Respect privacy and confidentiality	✓	✓	✓	✓	✓			
4 Integrity	✓	✓	✓	✓	✓	✓	✓	✓
5 Accountability	✓		✓	✓	✓	✓	✓	✓
6 Transparency	✓	✓	✓		✓	✓	✓	✓
7 Stay informed		✓	✓			✓	✓	

Data Ethics Tenets are strongly and clearly aligned with ASA Ethical Guidelines for Statistical Practice

Source: Park & Tractenberg (2023)

Navigating tensions between the frameworks

Results :
SPD 1, 2 and 2A and DET

All DET align with SPD 1 Responsibilities 2 and 4 (credibility and trust) and SPD 2.1.4 (instrument function); 2.3.4 (disclosure avoidance). See also 2.7.2 (protection), and 2A.5 (transparency).

<u>Areas of Mixed Alignment</u>: SPD 1 (relevance) could conflict with all DET

Source: Park & Tractenberg (2023)

OMB SPD Guidelines	T1 Uphold laws and regulations	T2 Respect public	T3 Respect privacy and confidentiality	T4 Integrity	T5 Accountability	T6 Transparency	T7 Stay informed
1.1 Relevant	(~)	(~)	(~)	(~)	(~)	✓	(✓)
1.2 Credible and accurate	✓	✓	✓	✓	✓	✓	(✓)
1.3 Objective						✓	
1.4 Trust	✓	✓	✓	✓	✓	✓	(✓)
2.1.1 Written plan	✓			✓	✓	✓	(✓)
2.1.2 Survey design	✓			✓	✓		(✓)
2.1.3 Response rate design	✓	✓		✓	✓		(✓)
2.1.4 Functioning components	✓	✓	(✓)	✓	✓	✓	(✓)
2.2.1 Appropriate frame	✓	(✓)		✓	✓	✓	
2.2.2 Notify respondents	✓	✓	✓	✓	✓	✓	
2.2.3 Balance quality v burden	✓	(✓)		✓	✓		(✓)
2.3.1 Appropriate data edits	✓	(✓)		✓	✓	✓	(✓)
2.3.2 Nonresponse analysis	✓			✓	✓	✓	(✓)
2.3.3 Quality for other study	✓	✓	✓	✓	✓	✓	
2.3.4 Avoid disclosure	✓	✓	✓	✓	✓	✓	✓
2.3.5 Evaluate data quality	✓	(✓)		✓	✓	✓	✓
2.4.1 Use theory and methods	✓			✓	✓	✓	✓
2.5.1 Analysis plan	✓				✓	✓	
2.5.2 Good statistical practice	✓		(✓)		✓	✓	✓
2.6.1 Review dissemination	✓	(~)	✓	✓	✓	✓	
2.7.1 Equitable dissemination	✓	✓		✓	✓	✓	(✓)
2.7.2 Data protection	✓	✓	✓	✓	✓	✓	
2.7.3 Survey documentation	✓					✓	
2.7.4 Public use microdata	✓	✓	✓			✓	
2.A.1 Methodological plan	✓			✓	✓	✓	(✓)
2.A.2 Sample selection standard	✓	(✓)		✓	✓	✓	
2.A.3 Interview guide standard	(✓)				✓	✓	
2.A.4 Systematic analysis	✓						(✓)
2.A.5 Transparent analysis	✓	✓	✓	✓	✓	✓	
2.A.6 Final reports standard	✓				✓	✓	
2.A.7 Reporting results standard	✓	(~)	(~)		✓	✓	

OMB Data Ethics Tenets

17

Results Summary: DET and PNP

Areas of tension: Principle 1 (relevance) Practice 7 (collaboration), and Practice 9 (dissemination of products) must be accomplished in balance with DET

Source: Park & Tractenberg (2023)

OMB Data Ethics Tenets (2020): / NASEM Principles and Practices (2021):	T1 Uphold laws and regulations	T2 Respect public	T3 Respect privacy and confidentiality	T4 Integrity	T5 Accountability	T6 Transparency	T7 Stay informed
PRI1 Relevance	(~)	✓,~	(~)	(~)	(~)	(~)	
PRI2 Credibility	✓	✓	✓	✓	✓	✓	✓
PRI3 Trust	✓	✓	✓	✓	✓	✓	✓
PRI4 Independence	✓	✓	✓	✓	✓	✓	
PRI5: Improvement	(✓)		✓		(✓)		✓
PRA1: Mission	(✓)					✓	
PRA2: Authority	✓		(✓)	✓	✓	✓	
PRA3: Commitment	✓	✓	✓	✓	✓	✓	(✓)
PRA4: Staff	✓			✓	✓		✓
PRA5: Research	(✓)		(~)			(✓)	✓
PRA6: Evaluation	✓	✓			✓	✓	
PRA7: Collaboration	~	~	(~)	(~)	(~)	(~)	✓
PRA8: Respect	✓	✓	✓		✓	✓	(✓)
PRA9: Users' Needs	~	~	(~)	(~)	(~)	✓,~	(✓)
PRA10: Transparency	✓	✓	✓	✓	✓	✓	(✓)

18

Questions?