Understanding Machine-Readability in Modern Data Policy

This paper provides an overview of why machine-readability matters and asserts that public policy for disclosure modernization should specifically require that both the values that entities report and the compliance standards for reporting those values be structured as machine-readable data. It concludes by offer three policy recommendations [documented as goals in this StratML rendition]

Legislators and executive agency policy making ... should address the three recommendations ... establishing disclosure modernization not just as best practice but as the standard practice. We summarize the practical implications of these recommendations:

- Government agencies and other data consumers will be able to automate validation of data quality, the completeness of a reported data set, and the consistency of that data set. Further, they will be able to maintain effectiveness even as data volume and velocity increase.
- The machine-readable data will be detailed, high quality, and timely, available for both conventional analysis and the innovative analysis achieved by machine-learning and artificial intelligence platforms.
- It will build confidence in compliance and financial systems because with this data, fraud is caught more readily and earlier.
- People will be able to understand who gets funding and when, and how that funding is being utilized. This will improve government accountability and the performance of programs.
- Audits and data analysis will be democratized because vetted data is made available earlier, with greater scope, precision, and completeness.
- The data is a shared resource for citizens. This transparency encourages citizens to trust and engage their government. This trust is essential when a nation faces major events and crises (e.g., a pandemic, a natural disaster, a man-made disruption to infrastructure and/or society).
- Policymakers should recognize that disclosure modernization as described here supports the legitimacy and the proper functioning of compliance and financial systems, and most importantly, of governments themselves.

Contents

Vision........................................................................................................................................................... 3
Mission......................................................................................................................................................... 3
Values .......................................................................................................................................................... 4
RECOMMENDATION #1. Machine-Readability ............................................................................................ 5
  1.1. Laws, Regulations & Guidance ........................................................................................................ 5
RECOMMENDATION #2. Intentions ............................................................................................................. 6
  2.1. Role, Purpose & Scope ..................................................................................................................... 6
  2.2. Standards .......................................................................................................................................... 6
RECOMMENDATION #3. Combination & Analysis ...................................................................................... 7
  3.1. Obstacles .......................................................................................................................................... 7
  3.2. Standards .......................................................................................................................................... 7
Administrative Information.......................................................................................................................... 7

Page 1
Data Foundation (DF)

Description:

The Data Foundation is a non-profit organization that seeks to define an open future for our data, for a better government and society. The Data Foundation provides research, analysis, education, and programming that aims to improve the country’s data policies. As the country’s leading data policy think tank, the Data Foundation supports informed dialogue and thought leadership on emerging strategies and applications for accessing, using, and applying data in ways that are useful for decision-making.

Stakeholder(s):

Dean Ritz:
Author — Board Member, Data Foundation, Senior Director of Digital Reporting Strategy, Workiva

Workiva Inc.:
Workiva is a registered trademark of Workiva Inc.

Contributors:
The author wishes to thank Christina Scheuer for substantial drafting assistance; to Workiva for funding this paper; and to the following people who reviewed and commented on later drafts: Rafael Gonçalves, C. Christina Ho, Joseph Howell, Mike Starr, Richard Szabo, Catherine Tsai, and Carol Volpe.

Christina Scheuer
Rafael Gonçalves
C. Christina Ho
Joseph Howell
Mike Starr
Richard Szabo
Catherine Tsai
Carol Volpe

Vision
Trust in institutions improves

Mission
To enhance transparency and accountability in our society
Values

**Data**: Application of data standards in our society offers potential to ensure potential data users have access to information about what data means. With clear delineation of data meaning, researchers, analysts, and other data users can effectively use data to create value for the economy, policymakers, and society.

**Standards**: Data standards promote common understanding and agreement on access to information. When made machine-readable—a format that can be accessed and used by computers—the efficiency and effectiveness of data analysis approaches are enhanced.

**Reporting**: The application of data standards and machine-readable data-reporting processes, such as with regulatory or compliance reporting, is currently underway as disclosure modernization sweeps through industry, academic, and governmental entities.

**Confidence**: Disclosure modernization supports the proper functioning of compliance and financial systems, and most importantly, of governments themselves. The benefits and practical implications are vast, including that standardization builds confidence in compliance and financial systems which quickly and reliably detect fraud, errors, and other concerns.

**Transparency**: Improving public policies related to disclosure modernization can support efforts to enhance transparency and accountability in our society, ultimately including improving public trust in institutions.

**Accountability**
RECOMMENDATION #1. Machine-Readability

Require machine-readability.

**Stakeholder(s)**

Policymakers:

Policymakers should require machine-readability when possible.

Public policy for disclosure modernization should require both the reported compliance values and the compliance standards to be machine-readable. Machine-readable compliance standards enable technical innovations for automated reporting and data validation to reduce the compliance burden. Encoding domain knowledge as machine-readable data measurably improves the opportunities for analysis by both conventional and machine-learning methods, reducing the cost of governmental and stakeholder oversight while also increasing its effectiveness.

**1.1. Laws, Regulations & Guidance**

Incorporate in proposed bills, regulations, and guidance documents requirements for machine-readable data.

**Stakeholder(s):**

Policymakers:

Policymakers could implement this recommendation by incorporating language in proposed bills, regulations, and even guidance documents requiring machine-readability, similar to the standard applied in the OPEN Government Data Act.
RECOMMENDATION #2. Intentions

Clearly communicate intent in legislative and regulatory actions.

Stakeholder(s)

Policymakers:
Policymakers should clearly communicate intent in legislative and regulatory actions on the role, purpose, scope of detail rendered as data, and applicability of data standards.

Legislators:
Legislators can specifically support implementation of this recommendation by including expectations and purposes for data standards when drafting legislation, through bill text or committee reports.

Regulators:
Similarly, regulators can better support this recommendation by clearly establishing expectations for standard effectiveness through proposed regulatory actions and guidance documents.

The second recommendation is a necessary technical complement to the first. We recommend that federal data policy pursue disclosure modernization where data standards are expressed as L4 taxonomies or L5 ontologies. The transition to disclosure modernization recognizes modernization as a necessary commitment by the regulated and the regulators to ensure the integrity of compliance, evidence-based policymaking, and capital markets. Modern data practices reflect the understanding that all data are not created equal. More sophisticated demands upon data require more expressive levels of agreement for data standards. Modern data policy should communicate intent by describing the levels of agreement for data that should be met by those implementing policy. For compliance reporting, the very real necessity for data effectiveness (as noted in Recommendation #1) is met only by the more expressive levels of agreement on data standards: L4 taxonomies and L5 ontologies.

2.1. Role, Purpose & Scope
Communicate the role, purpose, scope of detail rendered as data.

2.2. Standards
Communicate the applicability of data standards.
RECOMMENDATION #3. Combination & Analysis

Enable combination and analysis of data.

Stakeholder(s)

Policymakers:

Policymakers should encourage the adoption and use of open, consensus standards to encourage cooperation, efficiency, and innovation when drafting new data policies.

The goal is to enable agencies and data users—both public and private—to combine and analyze data (within statutory limits).

3.1. Obstacles

Minimize the technical and intellectual property obstacles to sharing and aggregating data.

The third recommendation is that technology choices should minimize the technical and intellectual property obstacles to sharing and aggregating data.

3.2. Standards

Recognize in legislative and regulatory actions the availability and prioritization of existing open standards.

Open source software and mature data encoding standards that are free to use without restrictions and maintained by voluntary consensus bodies are multipliers of innovation. Governments should follow private enterprise in realizing the benefits of this technical and social cooperation.

Stakeholder(s):

Policymakers:

Policymakers can specifically incorporate these approaches by explicitly recognizing in legislative and regulatory actions the availability and prioritization of existing open standards when implementing new directives and policies.