



October 21, 2024

1345 Avenue of the Americas  
27<sup>th</sup> Floor  
New York, NY 10105  
Phone: (202) 448-1985  
Fax: (866) 516-6923

Ann E. Misback, Secretary  
**Board of Governors of the Federal Reserve System**  
20th Street and Constitution Avenue NW  
Washington, DC 20551

Melane Conyers-Ausbrooks Secretary of  
the Board  
**National Credit Union Administration**  
1775 Duke Street  
Alexandria Virginia 22314-3428

Christopher Kirkpatrick, Secretary of the  
Commission  
**Commodity Futures Trading Commission**  
Three Lafayette Centre, 1155 21st Street  
NW, Washington, DC 20581

Chief Counsel's Office  
Attention: Comment Processing  
**Office of the Comptroller of the Currency**  
400 7th Street SW  
Suite 3E-218  
Washington DC 20219

Comment Intake—FDTA-INTERAGENCY  
RULE  
c/o Legal Division Docket Manager  
**Consumer Financial Protection Bureau**  
1700 G Street NW  
Washington, DC 20552

Vanessa A. Countryman  
Secretary  
**Securities and Exchange Commission**  
100 F Street NE  
Washington DC 20549-1090

James P. Sheesley,  
Assistant Executive Secretary  
Attention: Comments/Legal OES (RIN 3064-  
AF96)  
**Federal Deposit Insurance Corporation**  
550 17th Street NW  
Washington, DC 20429

Chief Counsel's Office  
Attention: Comment Processing Office of  
Financial Research  
**Department of the Treasury**  
717 14th Street NW  
Washington DC 20220

Clinton Jones, General Counsel  
Attention: Comments/RIN 2590-AB38  
**Federal Housing Finance Agency**  
400 Seventh Street SW  
Washington, DC 20219

RE: Proposed Rule: Financial Data Transparency Act Joint Data Standards  
OCC Docket ID OCC-2024-0012  
FRB Docket No. R-1837; RIN 7100-AG-79  
FDIC RIN 3064-AF96  
NCUA RIN 3133-AF57  
FHFA RIN 2590-AB38  
CFTC RIN 3038-AF43  
SEC File No. S7-2024-05  
CFPB Docket No. CFPB-2024-0034

Dear Agencies:

We appreciate the opportunity to provide input to the Agencies Financial Data Transparency Act (FDTA) rule proposal. We strongly support the legislation as a means to help Agencies provide actionable, good quality data to citizens, businesses, investors, policy setters, governments, and researchers. We support the adoption of the Legal Entity Identifier (LEI), and we support certain elements of the proposal as it relates to data standardization.

We have significant concerns however, about aspects of the proposal which will result in Agency collections of data that are not interoperable, will not reap the benefits of economies of scale that data standards can bring, and will incur unnecessary costs across the data ecosystem.

XBRL US is a nonprofit standards organization, with a mission to improve the efficiency and quality of reporting in the U.S. by promoting the adoption of business reporting standards. XBRL US is a jurisdiction of XBRL International, the nonprofit consortium responsible for developing and maintaining the technical specification for XBRL (eXtensible Business Reporting Language). XBRL is a free and open data standard widely used in the United States, and in over 200 implementations worldwide<sup>1</sup>, for reporting by public and private companies, as well as government agencies.

This letter provides our input to the rule proposal for the FDTA and includes as an Attachment responses to questions raised by the Securities and Exchange Commission (SEC) which may be of interest to all Agencies.

---

<sup>1</sup> See XBRL Project Directory: <https://www.xbrl.org/the-standard/why/xbrl-project-directory/>

## Executive Summary

The cost of government regulation is increasingly seen as a burden by many in Congress and the courts who look to limit the scope of regulatory agencies. At the same time, data collected by regulatory agencies is becoming more important for the functioning of a modern economy. With increases in geopolitical, environmental, market, liquidity, counterparty, and political risk, the need for timely and comprehensive information is essential to navigate these risks. The cost and burden of data collection must be reduced, and the accessibility of data must be improved. Both can be achieved through standardization.

Regulators may be reluctant to embrace standardization as they believe it limits their flexibility to regulate as they see fit. This is a trap that must be avoided. Digital standardization enforces a disciplined and structured approach that results in a regulatory framework that is transparent, robust, and unbiased. As regulators consider implementation of the FDTA, it is important to keep in mind what constitutes success: better data, reduced cost, increased flexibility. We should not settle for anything less.

To that end, we urge the Agencies to consider our recommendations, highlighted in this summary and explained in greater detail below. First, **we strongly recommend that Agencies adopt a single semantic data model structure (XBRL) rather than the properties-based approach described in the rule proposal.** The proposed approach will result in Agencies continuing to manage discrete, siloed datasets as they do today, that are not interoperable and cannot be commingled or automatically shared and inventoried.

Adopting a single semantic data model, as we explain in this comment letter, will help Agencies realize economies of scale, and will reduce costs for regulators and reporting entities, as well as users of the data: citizens, investors, policy setters, and researchers. The XBRL standard can express many types of data including financial and non-financial data; and can harmonize with other standards that uniquely express certain types of data such as the ACTUS standard for financial contracts, messaging standards such as FpML, FIX Protocol, ISO 20022; entity identifiers like the LEI; financial instrument classifications like CFI, FIBO, and UPI; and more.

Benefits of the single semantic data model structure include:

- Data produced by Agency collections will be machine-readable and ***machine-understandable***, eliminating the need for manual data entry and vetting, establishing a common digital language for all stakeholders. Data will be interoperable, shareable and can be commingled and inventoried together. This is feasible because even though the Agency, the reporting entity, and the data reported are *different*, the structure of the data is the *same*. The structured, granular, consistent nature of reported data lends itself to artificial intelligence and machine learning applications, which is becoming more critical for regulators, businesses and researchers handling high-volume information.
- **Data quality enhancements** because validation (business) rules can perform complex checking of accounting and regulatory standards, completeness, reasonableness checks and more. Validation rules created for one data collection can often be easily repurposed for other data collections because of the highly structured nature of the data.

- **Economies of scale** that reduce regulatory cost to collect and analyze data because all nine Agencies can leverage the same tools and databases. They do not need to build custom applications and can “borrow” and share tools from other Agencies, while still maintaining the uniqueness of their own data collection requirements and analysis. The same economies of scale that benefit government regulators will also reduce costs for reporting entities and data consumers. Software providers with applications that support one XBRL-reporting program often leverage the same applications for other XBRL-reporting programs. Development costs can be shared across many reporting entities, resulting in lower costs for those reporting. The cost of maintaining three separate products for three reporting situations will be higher than if a single application can be developed and costs shared across many. The same holds for analytical tools. Products that can be used for multiple datasets will be less expensive than products that must be tailored for single datasets and those savings will be passed on to reporting entities and data users.
- **Automation of Agency collaboration.** Following the same semantic data model structure also automatically coordinates the work of the Agencies, without the need to establish bureaucratic steering committees to monitor work and ensure collaboration.
- **Gives Agencies flexibility** to choose from multiple data transmission formats to “transport” their data including CSV, JSON, XML, and XHTML. The standards program will also be set up to adapt to new data transmission formats that may be introduced in future years, because the transmission process is separate from the semantic data model. Agencies can choose the transmission that is the best fit for their data.
- **Agencies will be able to update/revise reporting requirements more easily** and in a less costly fashion than the manual, paper-based process followed today; data preparers and data users will be able to adapt to updated requirements with minimal disruption. Time series data will remain intact even when reporting needs change.
- **Reporting requirements will be kept current** and have the flexibility to meet Agency and standard setter needs. Issuers will rely on the accounting standards they use today, for example, IFRS, GASB pronouncements, and FASB pronouncements, which will be kept current through taxonomies developed and maintained by the standards organization themselves. Agencies will be able to require Agency-specific reporting as well through taxonomies they develop that can be used seamlessly by issuers at the same time as taxonomies created by standards organizations.

We also recommend:

- **Require a taxonomy/schema for all Agency collections.** Even collecting “name” and “address” on an application has a schema with associated definitions, and relationships.
- **Re-use existing taxonomies where possible** such as the FASB GAAP Taxonomy used by public companies. Reporting entities such as banks and credit unions also adhere to FASB pronouncements. Recreating what is already available is costly, unnecessary and will result in data that is not interoperable.
- **Establish a governance framework for the Agencies** to facilitate continued sharing of information and standards development.

Furthermore, we urge the Agencies to eliminate potential off-ramps in the final rule that could derail the implementation such as the use of the term “to the extent practicable” which could be an easy exit when there is any pushback from stakeholders. Successful standards programs that yield the benefits outlined above require change and collaboration which is often difficult to embrace, for reporting entities, data users, and for the Agencies themselves. There are ways to ease the path for all stakeholders which could include compliance phase-ins or adopting creative pricing models.

Below are detailed explanations for each recommendation.

### **Require a single semantic data model structure (XBRL) to ensure interoperability, supported by multiple data formats to allow flexibility for Agencies.**

We agree with the proposed rule that Agencies should have the flexibility to choose the data transmission format most appropriate for their data collections; and we agree that standards should evolve as new technologies are introduced.

We do not however, support the proposed “properties-based approach” which allows Agencies to select a data transmission format based on four criteria<sup>2</sup>.

#### ***Why the properties-based approach is problematic.***

This method is grounded on the assumption in the proposal, that “...*data transmission or schema and taxonomy formats that have these properties are likely to be interoperable with each other*”<sup>3</sup>. This assumption is incorrect.

“A high degree of interoperability, so regulators and financial firms can analyze the same data using different systems” is named as one of the three key ingredients for the level of quality needed for financial stability in the first Office of Financial Research (OFR) Financial Stability Report<sup>4</sup>. The FDTA proposal for a properties-based approach does not address this critical component and is not even named as one of the required properties.

The general nature of the four properties will result in Agencies continuing to manage discrete, siloed datasets (replicating what we already have today) that cannot be commingled, inventoried, or shared. The result is that data will be generated that is *not interoperable*, drastically limiting the

---

<sup>2</sup> See <https://www.federalregister.gov/d/2024-18415/p-169> - (1) Render data fully searchable and machine-readable; (2) Enable high quality data through schemas, with accompanying metadata documented in machine-readable taxonomy or ontology models, which clearly define the semantic meaning of the data, as defined by the underlying regulatory information collection requirements, as appropriate; (3) Ensure that a data element or data asset that exists to satisfy an underlying regulatory information collection requirement be consistently identified as such in associated machine-readable metadata; and (4) Be nonproprietary or available under an open license.

<sup>3</sup> See <https://www.federalregister.gov/d/2024-18415/p-176>

<sup>4</sup> 2015 OFR Financial Stability Report: [https://www.financialresearch.gov/financial-stability-reports/files/OFR\\_2015-Financial-Stability-Report\\_12-15-2015.pdf](https://www.financialresearch.gov/financial-stability-reports/files/OFR_2015-Financial-Stability-Report_12-15-2015.pdf)



ability of governments and businesses to reduce costs and gain access to consistent, comparable data.

The illustrative example below addresses reporting by municipal bond issuances; however, it could apply to data collections for any of the nine Agencies.

Most government securities issuers follow the Governmental Accounting Standards Board (GASB) pronouncements. Just as publicly traded companies are required to follow the Financial Accounting Standards Board (FASB) pronouncements and to report financial statement data by using an XBRL Taxonomy maintained by the FASB, a logical step would be for the SEC to work with the GASB to develop a Taxonomy for state and local government issuers. The GASB has already begun the process of building an XBRL Taxonomy that could be used by governments to report their Annual Financial Reports (AFRs) and Annual Comprehensive Financial Reports (ACFRs). The GASB recently presented their own taxonomy development work at the July 29-30 conference, GovFin 2024: Municipal Reporting Workshop<sup>5</sup>.

If the SEC, however, opts for a custom XML or JSON schema, or a CSV or PDF/A implementation for government reporting, which would be consistent with the “properties-based” approach, the end result will be:

- Governmental issuers will face higher costs, because reporting applications will need to be custom-built by the same filing agents and software providers that already serve SEC-reporting entities.
- The SEC will incur the cost of building a custom schema to support the complexity of financial data (which XBRL was designed to do). The SEC will not be able to leverage the tools it uses today to database, extract, and analyze public company and investment management data; and it will not be able to use the same validation rules to check for data integrity. All of these assets will need to be custom-built to fit the custom governmental reporting schema.
- Data consumers, including the public, investors, researchers, academics, and data analytics providers, will also need to build or buy custom extraction, database, and analytics tools to fit the design of the custom model.

An attempt to achieve some level of interoperability between datasets built using different schemas would require detailed mapping to align how data model 1 exchanges data characteristics like time period, units of measure, data type, etc., with how data model 2 exchanges the same information.

This “mapping” process would need to be replicated over and over, by the collecting Agency, and by every data user.

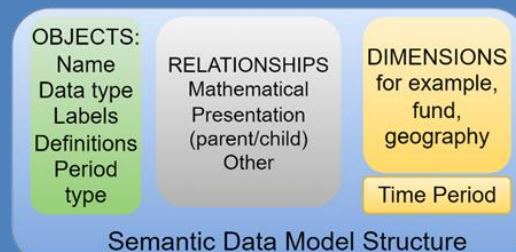
---

<sup>5</sup> Video from GovFin 2024: Municipal Reporting Workshop, July 30, 2024, Getting Started: Government Reporting Data Standards: [https://www.youtube.com/watch?v=km\\_wiFpyW8M&list=PLEyg0bSY11l4aecArcohh205443w60rcb](https://www.youtube.com/watch?v=km_wiFpyW8M&list=PLEyg0bSY11l4aecArcohh205443w60rcb)

Furthermore, when there is a change in reporting requirements - new data is required, definitions change, accounting standards change - the schema will need to be re-mapped across every Agency reporting application and data user. Every tool used to report, collect, or extract data will need to be revised to work with the new schema. In addition to the cost incurred for any change, the opportunity to introduce errors and inconsistencies across multiple entities and mapping procedures is enormous.

The swaps market is a cautionary tale about expecting interoperability when data standards are not clearly defined. The article, *The Data Reporting Challenge: U.S. Swap Data Reporting and Financial Market Structure*<sup>6</sup>, authored by Richard B. Berner, Robin Doyle, and Kenneth Lamar, provides a useful case study. The Dodd-Frank mandate for the collection of swaps data resulted in three new Commodity Futures Trading Commission (CFTC) rules for data collection. Rule makers, however, did not provide prescriptive rules on the details that needed to be reported, with the end result that, *“Left to using their own data schema, SDRs each developed their own reporting template, making comparability near impossible,*

## WHAT IS A SEMANTIC DATA MODEL STRUCTURE?



The term “single semantic data model structure” refers to a digital structure that expresses the features of data and the relationships that data have with other related data, in a consistent manner across data collections. It is a container that holds all the information needed to digitally understand the meaning of information and it structures this information the same way regardless of the data reported. The data model structure provides the building blocks to unambiguously define the data to be reported: information including object name, (for example Assets), data type, human-readable label, definition, dimensional characteristics and units of measure, along with the relationships between reported facts like mathematical (accounting, for example Assets = Current Assets + Noncurrent Assets) and parent/child relationships. All information shown in the diagram must be communicated along with the datapoint in order for a computer to read it without ambiguity.

The semantic model structure can represent all kinds of data from all kinds of entities, from companies that report following Financial Accounting Standards Board (FASB) GAAP, to governments that report following Governmental Accounting Standards Board (GASB) GAAP to banks submitting simple applications. The data reported is different, but it can all be expressed using the same semantic data model structure.

<sup>6</sup> The Data Reporting Challenge: U.S. Swap Data Reporting and Financial Market Structure: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3541248](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3541248)

*and sometimes requiring data elements that were of marginal use. As a result, reported data were often incorrect or late.”*

The CFTC and the SEC separately established U.S. swap data reporting collections that were inconsistent with each other, resulting in, as described by then-SEC Commissioner Scott D. O’Malia, ‘...*inconsistent reporting and variability in the data, as well as technology shortfalls combined with incongruent rules, have made the data presently unusable to the Commission.*’”

The authors note a key theme of the paper, “...*precise definitions, and data and technology standards are essential ingredients for implementing digital/technology solutions to improve the efficiency and accuracy of regulatory reporting,*” and conclude, “...*only by using uniform data standards will regulators be able to access and utilize the data made available. Standardization will have to be more advanced to meet regulatory requirements.*”

### ***Why a Single Semantic Data Model is preferable***

Standards like UPCs, QR codes, railroad track gauges and shipping containers, take an existing process or task, and improve its efficiency and effectiveness. One of the most important characteristics of any standard is compatibility.

Shipping container dimensions are standardized so that machines used to move them, such as top or side loaders or cranes, are also standardized; they can manage containers produced by any manufacturer and load them into any ship or truck. Without standards, custom loading machines would need to be built, and custom trucks or ships may need to be built in which to transport the containers. Economies of scale would not be achieved, costs would not be reduced, efficiencies would not be gained. This is what happens when there is a lack of interoperability in transport and exchange.

The degree of variation allowed in the draft FDTA rule proposal will result in the same lack of interoperability in the transport and exchange of data among Agencies. A properties-based data “standard” that lacks the level of specificity achieved by UPCs, QR codes, and shipping containers, will not be interoperable.

XBRL is a semantic data model structure that was developed to unambiguously capture financial and other types of business data. Reporting using the XBRL single semantic data model structure will enable interoperability even in a situation where the reporting entity relies on a different accounting standard or reports non-financial data such as greenhouse gas emissions or climate-related narratives. When a single semantic data model structure is used, filing agents and other software providers can adapt their existing tools to support all kinds of Agency data collections, which is much less costly than building custom applications for each reporting situation. Reduced product development and maintenance costs will come to issuers in the form of lower application and licensing fees.

When regulators opt for custom schemas, the result is data that is not interoperable as shown in the visual below.



#### Assets Represented in Custom Schema for Regulation Crowdfunding Companies

<totalAssetMostRecentFiscalYear>184886.00</totalAssetMostRecentFiscalYear>

This schema incorporates the time period of a reported value with the name of the fact, "assets"

#### Assets Represented in Custom Schema for Regulation A Companies

<totalAssets>73765284.00</totalAssets>

This schema captures the time period of the value in the report document

Custom-built XML schemas, like the two shown above, meet the four criteria of the properties-based approach. However, each was created separately even though the work was done within a single Agency, the Securities and Exchange Commission. The data generated by these two schemas is not interoperable and cannot be exchanged between computer systems. The schema on the top represents reporting for Regulation Crowdfunding companies. The schema on the bottom represents Regulation A companies. The use of custom schemas requires different tools to prepare, collect, and analyze data for crowd-funded companies and data for Reg A companies. Data is not interoperable despite the fact that the two schemas meet the four criteria defined in the FDTA proposal. Furthermore, public companies report using the XBRL standard with the schema defined by the FASB GAAP Taxonomy, for yet a third representation of the same data.

Filing agents and other providers that serve SEC filers often work with many types of SEC reporting companies. Under the current rules and the Agencies proposal, a provider that serves all three types of companies (Reg A, Reg CF, and public companies) must develop and maintain three separate products, which is more costly than supporting a single application for all. If a single schema were used for all reporting applications, the cost could be shared across companies of many types, enabling economies of scale that benefit reporting entities in the form of lowered costs from providers. This example is illustrative of how any Agency data collection may function under the rule as proposed.

When governments and public companies use XBRL Taxonomies, although the data reported follows different accounting standards, data produced by all types of reporting entities is interoperable, and can be shared, stored, inventoried, and commingled.

#### *How it supports any Agency data collection*

The image below shows how the single semantic data model structure represents the fact 25,399,809 which is the balance of cash and investments for governmental activities for the City of Soledad, California. The green bubbles represent the information about the fact in the XBRL taxonomy, which includes descriptive metadata and relationships that unambiguously define the fact. Concepts in this taxonomy are defined by GASB pronouncements which is noted in the green bubble on the lower left-hand corner of the diagram, "Reference = GASB Accounting Standards Codification, Section 150, Paragraph 535" as shown in the green bubble on the lower left-hand corner of the diagram. The yellow bubbles represent additional qualifying information for the fact that is contained in the report and prepared by the City of Soledad. All these attributes of the fact are defined in the single semantic data model structure.

# CITY OF SOLEDAD

## STATEMENT OF NET POSITION JUNE 30, 2023

Unit of measure = US dollars

Time Period = Period ending June 30, 2023

Reporting Entity = LEI for the City of Soledad

	Governmental Activities	Business-Type Activities	Total
<b>ASSETS</b>			
Cash and investments	\$ 25,399,809	\$ 16,374,182	\$ 41,773,991
Cash and investments with fiscal agent	-	457	457
Accounts receivable, net	378,628	2,335,490	2,714,118
Internal balances	(174,769)	174,769	-
Due from other governments	1,550,722	-	1,550,722
Interest receivable	11,102	6,153	17,255
Notes receivable	8,802,532	32,913	8,835,445
	-	2,109,968	2,109,968
	41,517,265	53,993,167	95,510,432
	77,485,289	75,027,099	152,512,388
<b>DEFERRED OUTFL</b>	3,736,651	593,179	4,329,830

Name =  
CashAndCashEquivalentsAndInvestments  
Label = Cash and Cash Equivalents and  
Investments

Period Type = instant

Definition = Cash and cash equivalents and  
investments includes not only current ...

Data Type = monetary

Reference = GASB Accounting Standards  
Codification, Section 150, Paragraph 535

### Relationships

Presentation Relationship:  
CashAndCashEquivalentsAndInvestments is  
a child to Assets

Calculation Relationship:  
CashAndCashEquivalentsAndInvestments  
positively (+) rolls up into Assets

The second image below depicts a fact reported by a public company, which adheres to FASB pronouncements. This data collection follows the same semantic data model structure as the fact shown above but it uses definitions, authoritative references, and other requirements of the FASB.

**Cantaloupe, Inc.**  
**Consolidated Balance Sheets**

Unit of measure = US dollars  
Time Period = Period ending June 30, 2024  
Reporting Entity = LEI for Cantaloupe, Inc.

As of June 30,

(\$ in thousands, except share data)	2024	2023
<b>Assets</b>		
<b>Current assets:</b>		
Cash and cash equivalents	\$ 58,920	\$ 50,927
Accounts receivable, net	43,848	30,162
Finance receivables, net	6,391	6,668
Inventory	40,791	31,872
	7,844	3,754
	157,794	123,383

**Name =**  
CashAndCashEquivalentsAtCarryingValue

**Label =** Cash and Cash Equivalents, at Carrying Value

**Period Type =** instant

**Definition =** Amount of currency on hand as well as demand deposits with banks or financial institutions, including ...

**Data Type =** monetary

**Reference =** FASB Accounting Standards Codification, Section S99, Subparagraph (SX 210.5-02(1))

**Relationships**

**Calculation Relationship:**  
CashAndCashEquivalentsAtCarryingValue positively (+) rolls up into Assets

**Presentation Relationship:**  
CashAndCashEquivalentsAtCarryingValue is a child to Assets

What is important to note from these two diagrams is that while the data and the accounting standards are different, the information about the facts (time period, labels, definitions, units of measure, the object concept, etc.) is all captured and transported the same way. For example, the units for both examples are in US dollars (USD) following the ISO 4217 currency standard and the time period for both follow the same ISO 8601 time period standard. Data prepared with the same structured approach can reside in the same database, and can be accessed using the same tools, even if the data is quite different. If these facts had the same associated time period, information could be queried on that date across both entities in the same data collection.

As another example of how data in a single semantic data model structure is interoperable, XBRL US maintains a single database<sup>7</sup> that houses information reported by: public companies to the SEC; public utilities to the FERC; European companies to the European Securities and Markets Authority (ESMA); and even government ACFRs prepared using the XBRL US ACFR Taxonomy<sup>8</sup>. The data is all structured identically regardless of accounting standard, reporting entity, or regulatory agency, so that it can be extracted and analyzed together.

Instead of the properties-based joint standard described in the proposed rule, the Agencies should require a single semantic data model structure that has the following characteristics: (1) is a structure that the Agencies can use for all FDITA-covered information collections; (2) is freely

<sup>7</sup> XBRL US Database of public filings (note that FERC, ESEF and SEC filings are publicly available; access to a limited ACFR dataset may be available on request): <https://xbrl.us/home/use/filings-database/>

<sup>8</sup> XBRL US Government Reporting Information Package (GRIP) Taxonomy: <https://xbrl.us/xbrl-taxonomy/2022-grip/>

available under an open license; (3) is software-agnostic; (4) consistently depicts attributes of financial data, including time, units of measure, data type, and other dimensional features; (5) has interchangeable components that can be reused (a modular approach will allow Agencies to share common concepts such as “Name” and “Address” which should have the same definition, name and characteristics regardless of collecting agency or entity); and (6) is able to easily handle the Legal Entity Identifier (LEI).

**When the same semantic data model structure is used for all Agency data collections, the result is consistently structured data that can be shared, stored, inventoried, commingled and analyzed together.**

Structured data that is interoperable is an optimal source for Artificial Intelligence applications as the consistent construction of each datapoint and its corresponding descriptive metadata will generate more useful outcomes from AI algorithms. According to a recent Department of Commerce (DOC) Request for Information (RFI), AI and Open Government Data

Assets Request for Information<sup>9</sup>, “*..today’s AI systems are fundamentally limited by their reliance on extensive, unstructured data stores, which depend on the underlying data rather than an ability to reason and make judgments based on comprehension.*”

The DOC RFI aimed to explore how to achieve better data integrity, accessibility, and quality because “*AI tools are increasingly used for data analysis and data access, so Commerce hopes to ensure that the data these tools consume is easily accessible and “machine understandable,” versus just “machine readable.”*” We agree that AI systems will be strongly supported by the availability of structured, standardized, interoperable datasets envisioned by the FDTA.

Each data set collected by an Agency is different, but the data is structured and defined should be the same. This approach will support Agency requirements of the Evidence-Based Policymaking Act of 2018<sup>10</sup>, and the Open Government Data Act, which establishes practices for the federal government to modernize its data management practices and improve efficiency to inform policy decisions. The single semantic data model establishes a long-term data collections program that can adapt to new technologies in the future.

### ***The role of data transmission formats***

The proposal names specific data transmission formats as being appropriate to fulfill FDTA requirements. The proposal names XBRL, XML, JSON, and XHTML (Inline XBRL) as being able to reference a schema; and PDF/A as being appropriate if the PDF format has advanced tagging features. We agree that the named formats can digitally convey information, however there are critical differences between the proposed methods as shown on the table in Appendix A. The visual below shows how XBRL harmonizes with various data transmission formats.

<sup>9</sup> See Department of Commerce Request for Information: <https://www.govinfo.gov/content/pkg/FR-2024-04-17/pdf/2024-08168.pdf>

<sup>10</sup> See <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>



The XBRL semantic data model structure has evolved over time to harmonize with various data transmission formats, starting with XML (in 2002), then adding XHTML (in 2011), followed by CSV (2021), and JSON (2021). This approach has proven successful at helping regulators worldwide adapt to new technologies, as needed, and as proven by the use of XBRL-XML, XBRL-CSV and Inline XBRL (XBRL-XHTML) versions by US regulators.



XBRL is a semantic data model structure which can be rendered as an XML file (XBRL-XML), an XHTML file (called Inline XBRL), as a CSV file (XBRL-CSV), or as a JSON file (XBRL-JSON). It was designed this way to ensure that it adapts to technological changes over time.

Using a single semantic data model structure like this gives Agencies the flexibility to select the data transmission format that best fits the data reported, for example CSV for high-volume repetitive datasets, and XHTML for datasets that are more useful when in both human and machine-readable form.

The SEC, for example, has data standards programs that collect data in XBRL-XHTML (Inline XBRL), and it also has a data collection program in XBRL-XML. The SEC does not however have XBRL-CSV implementations which we have recommended for several recent SEC programs. For example, the SEC final rule, Enhanced Reporting of Proxy Votes by Registered Management Investment Companies; Reporting of Executive Compensation Votes by Institutional Investment Managers<sup>11</sup>, requires Form N-PX to be prepared in custom XML rather than XBRL. The final rule states, “...with respect to XBRL-CSV, the Commission believes using the XBRL data model to define the elements and relationships featured in Form N-PX would add unnecessary complexity because Form N-PX consists of a relatively simple two-dimensional set of rows and columns, and does not feature any complex interlinking relationship among different rows. In addition, XBRL-CSV is not likely to create significant efficiencies in preparing and using managers' Form N-PX data because only a small number of managers are subject to a reporting requirement to file XBRL disclosures with the Commission.”

We disagree with the SEC's stated rationale in opting for custom XML. Even with a small number of filers as the SEC points out, software providers that already serve public companies and investment management companies that report in XBRL, could re-purpose existing tools for Form N-PX if it was prepared in XBRL; and data users that already extract data in XBRL format for

<sup>11</sup> See <https://public-inspection.federalregister.gov/2022-24292.pdf>



other purposes, could repurpose those tools for Form N-PX data extraction. With Form N-PX in custom XML, completely new tools for reporting and data extraction must be custom built. The greater expense incurred by software providers will be passed on to reporting entities, and to data users.

Furthermore, the SEC will incur additional costs with a custom XML approach. The Commission must build a custom XML schema, create custom tools to extract the data and maintain a separate database to store the data - all because Form N-PX data is not interoperable with other data that the Commission collects. Opting for XBRL-CSV rather than a custom XML schema would result in smaller file sizes; custom XML requires issuers to repeat identifying tags over and over, rather than referencing a taxonomy.

Other federal agencies have opted for varying XBRL approaches, picking the data transmission format that is the best fit for the data reported. The Federal Energy Regulatory Commission (FERC), for example, has a data standards program that collects financial statement data in XBRL-XML format and will soon introduce a data collection program for high-volume Electric Quarterly Report filing in XBRL-CSV.

The FERC Commission Order RM23-9-000 from October 19, 2023,<sup>12</sup> explains why they opted for XBRL-CSV, *“The Commission proposes to adopt a new EQR submission system based on the XBRL-CSV standard... XBRL-CSV applies the XBRL standard to the CSV format, the format favored by most Sellers [filers of the EQR]. The Commission believes that adopting the XBRL-CSV standard would preserve the efficiency and simplicity of CSV, while adding the flexibility associated with the XBRL standard...the Commission believes that transitioning the EQR system to the XBRL-CSV standard will make information easier for Sellers to submit and for data users to retrieve, while also decreasing the costs, over time, of preparing the necessary data for submission and complying with future changes to the Commission’s filing requirements.*

*One benefit of the proposed XBRL-CSV system is that it would allow Sellers to continue to prepare and review their data in Excel spreadsheet format and then submit their data in CSV format.... the proposed new system would allow Sellers to use Excel to prepare multiple, smaller transaction files, which filers could then save as CSV and submit multiple transaction files without needing to combine them into one large transaction file...Another benefit of the proposed XBRL-CSV system is that it would save Sellers time in preparing their filings by allowing them to check their EQR submission for most errors in real-time by using the publicly available FERC EQR taxonomies and related documents without first submitting files to the Commission.”*

We agree that data transmission formats are important to transport data, but on their own, they are not sufficient to generate data that is fully computer-readable, computer-understandable, and interoperable. Combining the semantic data model structure with a data transmission format provides digital, interoperable, understandable data.

---

<sup>12</sup> See [https://elibrary.ferc.gov/eLibrary/filelist?accession\\_number=20231019-3066](https://elibrary.ferc.gov/eLibrary/filelist?accession_number=20231019-3066)

### *Review of schema and taxonomy formats named in the proposal.*

While all the named data transmission formats are capable of digitally conveying information in a document, only XML, JSON and XBRL can reference a common data language (schema or a taxonomy) that can be reliably parsed and analyzed. PDF/A can contain and transport free-form information, but there is currently no specification that allows it to reliably reference a schema that would generate structured data. While this may be feasible at some point in time, the PDF/A specification currently does not have that capability today and should not be allowable to fulfill the FDTA until such time as it has this capability.

XML, JSON, and XBRL can transport metadata about facts reported by referencing a schema or taxonomy which conveys explanatory information that helps the receiver of the information understand what the fact represents (for example, "Who reported the data? What was the time period? What is the definition for the fact?")

However, this is where XBRL diverges from XML and JSON.

XML and JSON<sup>13</sup> have the flexibility to support a custom-built schema; "custom" means the data structure can be designed *in whatever way the developer wishes*. XBRL, however, only supports a rigidly structured, standardized XBRL Taxonomy. The rigidity of the XBRL structure means all data generated using an XBRL standard is structured and produced in a standardized fashion; it has a standardized inflexible structure to transport both the fact and the metadata that further explains what the fact represents.

The SEC chose a custom XML schema for its implementation of Schedules 13D/13G, noting in the final rule, Modernization of Beneficial Ownership Data<sup>14</sup>, "... 13D/G-specific XML is more suitable than XBRL for Schedules 13D and 13G because it facilitates the use of a fillable form that should result in a lower cost of complying with the structured data requirement compared to XBRL, particularly for smaller and infrequent filers. Under an XBRL requirement, filers (including smaller and infrequent filers) would incur costs and burdens associated with tagging the disclosures (e.g., software licensing costs, time spent applying tags) or with paying a third party to do so. Thus, although some Schedule 13D and Schedule 13G filers, such as those currently subject to Inline XBRL reporting requirements (e.g., filers that are Commission registrants) or that otherwise have experience with XBRL may realize some efficiencies under an XBRL alternative, we believe the cost savings expected to arise from having a fillable form option under the 13D/G-specific XML requirements would have a more substantial positive impact with respect to filers as a whole."

While we agree that access to a fillable form, particularly for infrequent filers, is optimal, the SEC could create a fillable form that generates data in structured XBRL format. There would be no additional resources required to create an XBRL taxonomy versus a custom XML schema, and the data produced could then be extracted, analyzed, and used, by the same tools that the SEC

---

<sup>13</sup> Only the JSON-LD (linking data) specification has this capability.

<sup>14</sup> See <https://www.sec.gov/files/rules/final/2023/33-11253.pdf>

uses today (and that investors and data aggregators use as well) with structured XBRL data. The economies of scale would be beneficial for the regulator and for other data consumers.

While the terms “rigid” and “inflexible” are often considered to have negative connotations, it is the opposite with standards. It would be a chaotic world if electricians had the latitude to install wiring and current however they chose and had no guidelines to follow. Or if UPC codes were not truly “universal” and every manufacturer had their own custom-made version (with custom-built scanners and custom-designed codes for their products). By definition, restrictions are what make standards, *standard*.

The importance of concrete structure in standards is shown in the Federal Energy Regulation Commission (FERC) program. The FERC, which initiated the upgrade of their data collection using custom XML, switched to XBRL instead; and made the following statements in a recent webinar<sup>15</sup>:

*“We chose it [XBRL] over custom XML because there were a lot of efficiencies to gain in adopting XBRL and a lot of cost savings. We relied on our larger agency partners, the SEC and the FDIC who already had XBRL collections in place, and we learned from their experience and made some different choices for our system which I think paid off in the end.”*

*“We started with a custom XML solution because that’s what we had done in the past with all of our other collections ... Honestly, the more I looked into XBRL, the more I talked to other experts and to the SEC, XBRL answered a lot of the questions already ... about how to deal with time on a field... tons of questions that we would have had to answer in a custom XML implementation were already answered with XBRL. It also provided the flexibility to allow us to add our own rules on top of it. So, it just pushed the project months forward by being able to adopt XBRL.”*

Separately, the Federal Deposit Insurance Corporation (FDIC) noted in a video<sup>16</sup> that they chose XBRL because it readily supports financial data, “... the FDIC didn’t want to ... create another ‘XML-flavored’ business reporting language. We didn’t need to do that because XBRL already had all the definitions and everything that we needed to implement a solution.”

Unlike JSON or XML, XBRL has a standardized mechanism to communicate units of measure (for example, acres, square miles, dollars, percent), time periods, level of precision, human-readable labels, and definitions. XBRL assigns each fact a predefined data type to explain if the reported value expresses an integer, a monetary value, or a text string, for example. XBRL explicitly identifies the reporting entity and associates that entity with every fact reported. XBRL has a standard method to represent accounting relationships, presentation relationships, and other types of relationships that may exist between facts.

If an Agency opts for XML or JSON, it must create from scratch the structure to represent these relationships and metadata. Creating a custom schema for data that can be presented in XBRL

---

<sup>15</sup> XBRL US Webinar, September 11, 2024, FERC: How Regulators Use Technology to Improve Efficiency, <https://xbrl.us/events/240911/>

<sup>16</sup> Video: Government Use of Data Standards - Conversation with the FDIC: <https://xbrl.us/news/regulator-video/>

is recreating the structure that XBRL already has. More importantly, if an Agency creates a structure based on a specific data collection, it is *highly unlikely* that it will build one that is interoperable with other custom XML or JSON schemas at other Agencies.

Neither CSV nor PDF/A, the other possible formats named in the FDFTA rule, have any of the characteristics noted above, which are necessary for the unambiguous exchange of data.

In referencing PDF/A the rule proposal states, “*PDF may satisfy the standard if the data within the PDF conforms to specification “A” (PDF/A) that uses advanced features for tagging fields with a reference schema and taxonomy and provides necessary metadata that allows for automated data extraction.*” The PDF/A specification does allow the inclusion of information along with a fact, but it cannot reference a schema or taxonomy. While, in theory, the specification could be upgraded, standards development takes time and is not currently viable given the FDFTA rollout timing. Furthermore, PDF/A files that reference a schema/taxonomy would still require “tagging” of individual facts within the PDF file and it would not be the same as the process followed today when preparing a PDF, of creating a report in another format like Excel or Word, and then transforming it with a click of a button to PDF.

Many reporting entities, including government bond issuers, nonprofits, and companies today submit PDF formatted financials and other disclosures to fulfill requirements. There has already been pushback from a segment of the government issuer community, and there will likely be concerns raised from others that report in PDF format and that wish to keep the status quo.

We recognize that some reporting entities are likely to have a learning curve to climb, however we are concerned that PDF/A was included as a possible solution to serve as an off-ramp to some data collections. Transitioning to PDF/A today, given the current state of the PDF/A specification, would result in no change in the usefulness of the data, or the ability of data users to gain access to machine-understandable, automatable data. There will be a cost in transitioning to PDF/A for all involved, from issuers to data users, but **no benefits**.

XBRL taxonomies, again because of their structure, are compatible and interchangeable. Taxonomy modules can be re-used and combined, a practice used extensively by the SEC today for reporting by public companies and investment management companies. In fact, when public companies report, they typically use multiple taxonomies - some developed by the FASB (for FASB GAAP financials), and some developed by the SEC (for document and entity information). This facility makes it easier to: ensure data interoperability, and update and revise reporting requirements as needed, by the regulator or standard setter responsible for a particular reporting requirement. XBRL taxonomies enable long-term compatibility of data by allowing changes to be made in reporting requirements without “breaking” time series.

Furthermore, the XBRL structure lends itself to standardized validations (business rules) which can be used to check accounting rules (for example, Current Assets + Noncurrent Assets must total to Assets), reasonableness, reporting completeness, and more. While validation rules can be built for custom XML and JSON schemas, they must be created “from scratch” for each



schema. Validation rules created for one taxonomy can often be re-used for a separate taxonomy with minimal revision because of the consistent structure of the data.

### *Experiences of Regulators Using a Single Semantic Data Model*

The Federal Deposit Insurance Corporation (FDIC) began collecting data from banks using a single semantic data model in 2005. Their implementation resulted in immediate efficiency gains for members of the Federal Financial Institutions Examination Council (FFIEC) including<sup>17</sup>:

- 95% of data received met FFIEC validation requirements
- 100% of data received met FFIEC mathematical validation requirements versus 70% in the legacy system
- Data was publicly available immediately after calendar quarter end, versus weeks later in the prior legacy system
- Staff productivity in handling bank data increased 10-33%
- Data could be distributed to end users at agencies within one hour versus within several days in the legacy system

When the FERC began researching how to modernize their data collection, they started with XML but switched to XBRL after identifying difficulties making an XML program work for their needs. The flexibility of XML meant they were effectively “starting from scratch” building schemas to represent financial and energy data; and given the custom nature of the schemas they were creating; they had to create custom applications to work with the data.

The final FERC rule<sup>18</sup> states, “...the XBRL standard includes all the advantages of the XML format, such as its non-proprietary nature, its efficient sharing of data across different information systems, ... while also structuring the data with tags that utilize standard taxonomies to capture the inherent characteristics of the information as well as the value of the data... the XBRL standard is required for filing forms by a number of other federal agencies... XBRL is an international standard that enables the reporting of comprehensive, consistent, interoperable data that allows industry and other data users to automate submission, extraction, and analysis... the use of XBRL would facilitate the implementation of changes to its reporting requirements by enabling future changes without the need for costly development procedures.”

The FERC data modernization project<sup>19</sup>, which included project management, development of an XBRL taxonomy and (business) validation rules, training for internal staff, upgrades to FERC’s electronic filing system, data consumption tools and APIs, conversion of 10 years of historical data from Visual FoxPro files, and development, testing and implementation of a deployment plan, cost \$6 million to date, and is estimated to total \$7.4 million spent over six years when complete. The XBRL taxonomy for FERC reporting has ten forms with 303 individual schedules.

---

<sup>17</sup> See FFIEC, Improved Business Process Through XBRL: A Use Case for Business Reporting: <https://xbml.us/wp-content/uploads/2007/12/20060202FFIECWhitePaper.pdf>

<sup>18</sup> FERC Final Rule: <https://www.federalregister.gov/documents/2019/06/27/2019-13588/revisions-to-the-filing-process-for-commission-forms>

<sup>19</sup> See USASPENDING.gov, Award Profile Contract Summary, Department of Energy: [https://www.usaspending.gov/award/CONT\\_AWD\\_89603019P0018\\_8960\\_-NONE\\_-NONE-](https://www.usaspending.gov/award/CONT_AWD_89603019P0018_8960_-NONE_-NONE-)



The FERC noted in a recent webinar<sup>20</sup>, *“We delivered this project on time and on budget... This was a glowing success for FERC because it didn’t take more time and wasn’t much more expensive than what was considered in the original spend.”*

The FERC also called out the ease of making changes when reporting needs change, *“It also allows us to change things easier. When reporting requirements change, instructions or validation rules need to change, we have internal software applications that we use that can affect the taxonomy, publish a new version, change the validation rules, and schedule them for when they need to go out.”*

The single semantic data model approach has a 19-year history of successful use by the FDIC, the SEC, and the FERC in the US; and is in extensive use in non-US markets. The single semantic data model structure has been adopted in countries including Australia<sup>21</sup> and the Netherlands<sup>22</sup> where it is called Standard Business Reporting. The Netherlands estimates that they have been able to reduce administrative burden on regulators and business by 25% in five years with the XBRL standard estimated to account for €750 million of the initial savings.<sup>23</sup>

The United Kingdom’s HMRC Companies House<sup>24</sup> alone requires over 4.5 million companies<sup>25</sup> to report in structured (XBRL) format. There are 200+ global regulatory implementations<sup>26</sup>, ranging from financial regulators to tax authorities to ESG reporting.

## **Eliminate off-ramps that will dramatically limit the efficacy of the program.**

We urge the Agencies to remove the term *“to the extent practicable”* under the section on properties of a data standard which appears on page 71. This clause eliminates all responsibility of the Agencies to adopt any of the requirements of the FDTA and may result in Agencies selectively choosing which data collections will have data standards requirements and which will not. The implementation of standards requires change and there is likely to be pushback from reporting entities and others who may not see immediate benefits which are usually recognized over time. The public good of a robust data standard should outweigh an interest in preserving the status quo, if the government’s data is to be useful among agencies and by all Americans.

---

<sup>20</sup> How Regulators Use Technology to Improve Efficiency: <https://xbrl.us/events/240911/>

<sup>21</sup> SBR in Australia: <https://www.sbr.gov.au/>

<sup>22</sup> SBR in the Netherlands: <https://business.gov.nl/regulation/standard-business-reporting/>

<sup>23</sup> Standard Business Reporting Case Study: Tangible benefits of data standards for business and regulators: <https://xbrl.us/wp-content/uploads/2023/07/SBR-Case-Study.pdf>

<sup>24</sup> UK HM Revenue & Customs, XBRL guide for businesses: <https://www.gov.uk/government/publications/xbrl-guide-for-uk-businesses/xbrl-guide-for-uk-businesses#introduction>

<sup>25</sup> UK Companies House Official Statistics, Incorporated companies in the UK April to June 2024: <https://www.gov.uk/government/statistics/incorporated-companies-in-the-uk-april-to-june-2024/incorporated-companies-in-the-uk-april-to-june-2024>

<sup>26</sup> XBRL Project Directory: <https://www.xbrl.org/the-standard/why/xbrl-project-directory/>

As noted in Marc Levinson's book, *The Box, How the Shipping Container Made the World Smaller and The World Economy Bigger*<sup>27</sup>, "Only with time, as container shipping developed into an entirely new system of moving goods by land and sea, did it begin to affect trade patterns and industrial location. Not until firms learned to take advantage of the opportunities the container created did it change the world. Once the world began to change, it changed very rapidly; the more organizations that adopted the container, the more costs fell, and the cheaper and more ubiquitous container transportation became."

## **Require a taxonomy/schema for every data collection.**

According to the proposed rule, "Not all Agency collections of information have a schema and taxonomy associated with them, as a schema and taxonomy may not be appropriate."

Any data collection that will be inventoried, posted to a database, or shared, should have a supporting schema/taxonomy. This is necessary to define the logical configuration of the data being reported and allow the reporting entity, as well as the data collector and user to understand how the data is organized and structured. The taxonomy establishes a common (agreed upon) language for what is reported. Without a taxonomy, a computer cannot understand what a fact represents or how it relates to other data; data will not be prepared or collected within the constraints of having uniform definition and structure; reporting will generate data that is ambiguous, requires manual review, and cannot be automatically consumed and understood by computers.

Even if only two facts are reported on a form, for example, Corporate Title and Address as shown on the form below,<sup>28</sup> a schema is necessary to enable linking these facts across different data collections and defining what they represent in a consistent manner.

---

<sup>27</sup> Marc Levinson, *The Box, How the Shipping Container Made the World Smaller and the World Economy Bigger*, Second Edition, Princeton and Oxford, Princeton University Press, 2016, page 17.

<sup>28</sup> For example, Application to the Board of Governors of the Federal Reserve System for Membership in the Federal Reserve System - FR 2083: <https://www.federalreserve.gov/apps/reportingforms/Download/DownloadAttachment?guid=f2104d38-4257-4bb2-9458-ab388400476b>



## Application to the Board of Governors of the Federal Reserve System for Membership in the Federal Reserve System—FR 2083

An application is authorized by Sections 9 and 19 of the Federal Reserve Act, 12 U.S.C. §§ 321-328; and related provisions of law, and Section 208.3 of Regulation H.

NOTE: A bank located outside the United States in a dependency or insular possession, which desires to be admitted to the System under Section 19 of the Federal Reserve Act (12 U.S.C. § 466) should contact the Federal Reserve Bank of which it wishes to become a member or at which it wishes to maintain reserves.

Date of Application:   
Month / Day / Year

Corporate Title of Applicant

Proposed Corporate Title of Applicant, if applicable

Street Address of Head Office

City

State

Zip Code

County

Most data collected by the Agencies (as found in the OFR Interagency Data Inventory<sup>29</sup>) represent financial statements or non-financial data or may express inputs to a defined form or responses to survey questions. All collections include data that should be represented by a semantic data model structure that includes relationships between the reported facts so they can be understood by computers.

Financial statements have dimensional characteristics such as data broken down by fund, geography, or time period. Non-financial data like the FDIC Interagency Appraisal Complaint Form<sup>30</sup> categorizes complaints received by topical issue, location, and type of property. Capturing dimensional characteristics in digital format allows regulators and others to easily analyze trends. This is accomplished by establishing a taxonomy.

We urge the Agencies to create digital schemas for **all** data collections that are shared, inventoried, or maintained in a database, as all of these data collections will already have some form of schema, either in digital or paper formats. In addition, we maintain that most Agency data collections should reference a taxonomy to ensure the greatest efficiency and consistency of data. The use of schemas and taxonomies not only helps in data collection but in ongoing data management and provides a comprehensive inventory of all data reported. A taxonomy gives regulators the ability to revise and update reporting requirements and instructions once (in the taxonomy); changes are automatically communicated to preparers and data users simultaneously, improving communication, efficiency, and consistency across the supply chain.

If there are situations where a schema is not required, we urge the Agencies to unambiguously define those situations in the final rule to eliminate confusion and ensure that Agencies do not shortcut the approach to standards development.

<sup>29</sup> OFR Interagency Data Inventory: <https://www.financialresearch.gov/data/interagency-data-inventory/>

<sup>30</sup> FDIC Interagency Appraisal Complaint Form: <https://omb.report/icr/202206-3064-001/doc/122261400>

## Re-use existing taxonomies where possible.

Good quality, interoperable data is only possible when data is reported using a consistent, common language, and when data is structured in the same way. There are many taxonomies already in use that should be re-used for entities that report using the same information model. For example, the US GAAP Taxonomy, maintained by the FASB should be used by every entity that reports using FASB pronouncements, regardless of type of entity or regulator to which they report.

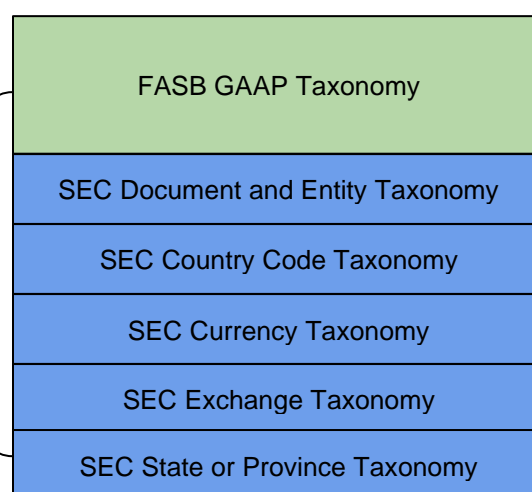
Every reporting entity that follows FASB pronouncements, which includes banks, credit unions, and public companies, should prepare their financials using the US GAAP Taxonomy. This method will eliminate duplicating the work to develop taxonomies; it will ensure that reported data is interoperable, reducing workload on reporting entities and ensuring the creation of consistent, comparable data.

State and local governments that report under GASB pronouncements need to follow the Governmental Accounting Standards Board (GASB). The SEC should collaborate with the GASB as they did in 2009 with the Financial Accounting Standards Board (FASB) on the development of the US GAAP Taxonomy for public companies. Governments that follow the GASB GAAP accounting standard should be required to follow a GAAP Taxonomy developed and maintained by the GASB.

Agencies that leverage a taxonomy built and maintained by a standard setter like IFRS, GASB, or FASB, can also require the reporting of additional facts beyond what the accounting standard setter includes, for example document and entity information.

As noted earlier, taxonomies are modular. Regulators can require reporting entities to use concepts from more than one taxonomy. The SEC for example, maintains its own taxonomies which are used by public companies at the same time they are using the FASB GAAP Taxonomy. Software tools present the issuer with all the elements needed from all required taxonomies in a seamless process. The issuer goes into the set of taxonomies as shown at right at a single-entry point

Reporting  
Entry Point



that gives access to all the concepts needed to prepare their financial statements. The SEC maintains separate smaller taxonomies that represent concepts needed for the document and the entity, such as identifiers, country, currency, exchange, and state or province. These SEC-maintained taxonomies operate in concert with the larger FASB GAAP Taxonomy. Similarly,

companies that report under IFRS accounting standards use the IFRS Taxonomy along with certain SEC-specific taxonomies to prepare their submission to the SEC.

The ability to create “entry points” that access multiple taxonomies allows multiple Agencies to use the same taxonomy for a particular accounting standard and retain the ability to tailor certain features for Agency-specific reporting needs that may not be covered by the accounting standard.

## **Build taxonomies in adherence to the single semantic data model (XBRL) structure.**

Taxonomies must be built following common principles across Agencies. By adhering to a single semantic model structure, Agencies will follow a consistent structure in defining reporting terms, required metadata, and relationships between the terms. The consistency of taxonomies will ensure that data can be reported, collected, extracted, and analyzed in the same way, which will ensure the lowest possible costs across the reporting ecosystem.

Because taxonomies are modular, we also support creating a “base taxonomy” which could contain elements that are used by multiple Agencies, such as “Company Name” and “Company Address.” All Agencies could leverage the base taxonomy and expand upon it by building their own taxonomies to include elements needed for agency-specific data collections. This approach was described in detail in the white paper, Data Standards, and the Financial Data Transparency Act (FDTA)<sup>31</sup>

One of the many benefits of using a single semantic data model structure is that Agencies can work on their own to build taxonomies to express their own data collections without the need to coordinate efforts and reach mutual agreement with other Agencies. If Agencies adhere to the single semantic data model structure, their taxonomies (and the data generated using those taxonomies) will be automatically interoperable and shareable. This approach can eliminate a significant amount of bureaucracy.

This approach can be further illustrated by our electrician’s example. An electrician installing wiring and current transformers at one house does not need to check with the electrician down the street performing the same job on a different house. Both follow the local electrical codes and standards on their own with the result that each home is set to use electricity in the same way. Different homes, different electrical needs, but both can use electricity to meet homeowner needs.

Collaboration between Agencies also enables the sharing of software tools, database structure, and validation rules which enables economies of scale and reduces costs across the federal government.

---

<sup>31</sup> Data Standards and the Financial Data Transparency Act (FDTA): <https://xbrl.us/wp-content/uploads/2022/04/data-standards-fdta-2023.pdf>



## Require the LEI as proposed.

We support the Legal Entity Identifier (LEI) as proposed. The XBRL specification supports the LEI and is already included in SEC and FERC taxonomies. To accommodate municipal bond issuance, standards need to be established to define the relationships that governments have with other government entities, for example, component units, obligors, and funds. XBRL can be used to support this because of the relationship linking features of the XBRL specification. This is described in greater detail in the paper, *Identifying the Obligor for Municipal Securities*<sup>32</sup>.

Furthermore, we encourage the SEC to consider the importance of educating government entities to help them understand what the LEI is, and the difference between the LEI and other identifiers with which they are more familiar such as the UEI, EIN, CUSIP, and DUNS.

Management of the LEI needs to be clearly defined given the relationships between reporting entity governments, separately audited funds, departments, and agencies of the reporting entity, as well as blended and discretely presented component units. We recognize that these issues will be covered in greater detail in the individual agency rules that will come out in the second phase of the FDITA, however it is important to begin considering these issues as soon as possible.

We support other identifiers named in the proposal including ISO 8601, ISO 4217, standardized state codes and country codes. We also ask the Agencies to consider establishing the verifiable LEI (vLEI) for the automated authentication and verification of legal entities.<sup>33</sup> GLEIF demonstrated the use of the vLEI to sign its 2023 Annual Report: each of GLEIF's Chair, CEO, and auditors cryptographically signed the report's contents, so that, forever after, anyone consuming the report's contents, including the financial statement data, would be able to ascertain if the contents have been tampered with.<sup>34</sup> The XBRL specification supports the vLEI.

## Establish a governance structure to ensure Agencies work together going forward.

The proposal states that the Agencies “*expect to work together on the adoption of the established joint standards in the Agency-specific rulemakings or other Agency actions,*” however it includes no requirement or details establishing a governing process to work together. This general statement is too haphazard an approach to the coordination and collaboration required to promote interoperability of financial regulatory data across members of the FSOC. The experience of the SEC and the CFTC, each of which in the last decade initiated swap data reporting collections that were inconsistent with the other, exemplifies what can happen when Agencies fail to work together: Agencies collect data that they then find is “unusable.”<sup>35</sup> The factors that historically

---

<sup>32</sup> Identifying the Obligor for Municipal Securities, July 2024: <https://cache.xbrl.us/wp-content/uploads/2024/05/Identifying-the-Obligor-for-Municipal-Securities-July-2024.pdf>

<sup>33</sup> See “Introducing the verifiable LEI (vLEI),” GLEIF, <https://www.gleif.org/en/vlei/introducing-the-verifiable-lei-vlei>.

<sup>34</sup> See “GLEIF Annual Report,” GLEIF, <https://www.gleif.org/en/about/governance/annual-report>.

<sup>35</sup> See Richard B. Berner, Robin Doyle, and Kenneth Lamar, *The Data Reporting Challenge: U.S. Swap Data Reporting and Financial Market Infrastructure* (Nov. 2020) (n. 19, quoting then SEC Commissioner Scott D. O'Malia), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3541248](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3541248).

have impeded progress in data standardization for financial regulatory reporting are well-documented and should be studied as lessons learned.<sup>36</sup>

Agencies can coordinate efforts effectively through partnerships with standard setting organizations that reporting entities follow today. The GASB, FASB and the IFRS Foundation, as well as climate disclosure standards organizations such as the International Sustainability Standards Board (ISSB) (which is also part of the IFRS Foundation), come to mind as standards organizations that would be best equipped to build taxonomies to support many of the Agencies data collections. This approach will automatically identify those areas where Agencies may be able to “share” taxonomies, for example FASB GAAP is followed by public companies, credit unions, and banks. Agencies can coordinate taxonomy use, the creation of validation (business) rules to check data quality, as well as database and data extraction and analysis tools, to save time and money.

For those data collections that are strictly designed and used by a single Agency, Agencies can build their own taxonomies which provide highly structured, searchable inventories of collected data (following the single semantic data model structure). The structured nature of the taxonomies will allow for easy review and comparison of data collected across all Agencies.

We urge the Agencies to establish a plan for ongoing coordination to monitor outcomes, share ideas and approaches, and adapt to changes in technologies, industry, and reporting needs. A logical cross-agency forum would be the FSOC’s Data Committee, which was chartered, among other things, to “[facilitate] information sharing and coordination among member organizations on data related matters, including data standardization”<sup>37</sup>

The same goes for monitoring developments related to data standards. Inherent in the directive in Financial Stability Act Section 124(c)(1)(B)(v) to “*incorporate standards developed and maintained by voluntary standards bodies*” is an obligation of the Agencies to monitor on an ongoing basis the work of relevant voluntary standards bodies. The United States Treasury’s OFR would be well-positioned to perform this role and make recommendations to the implementing Agencies because its staff members have for many years actively participated in standards development work.

If the Agencies were to establish the single semantic data model structure for information collections, the process for implementing Agencies to incorporate XBRL in the Agency-specific rulemakings could proceed as follows.

---

<sup>36</sup> See Richard Berner and Kathryn Judge, “The Data Standardization Challenge,” in *Systemic Risk in the Financial Sector: Ten Years After the Great Crash*, edited by Douglas W. Arner, Emiliós Avgouleas, Danny Busch, and Steven L. Schwarcz (2019), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3323719](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3323719).

<sup>37</sup> See “Charter of the Data Committee of the Financial Stability Oversight Council” (last accessed Sep. 14, 2024), available at [https://home.treasury.gov/system/files/261/The%20Council%26%23039%3Bs%20Committee%20Charters\\_1.pdf](https://home.treasury.gov/system/files/261/The%20Council%26%23039%3Bs%20Committee%20Charters_1.pdf).

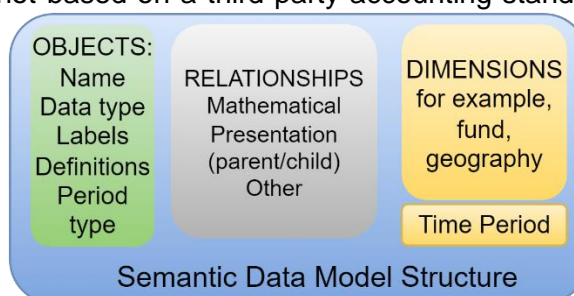
### ***Step 1. Identify organizations that set standards followed by Agency data collections.***

Agencies with reporting entities that follow statutory reporting standards such as FASB, GASB or IFRS, should coordinate with those standard setting organizations. The standards organization can build taxonomies to represent that reporting domain. Agencies with reporting entities that use the same accounting standard, for example, NCUA, the FDIC, and the SEC, should coordinate efforts to use the same taxonomy.

Agencies may collect additional data beyond what is included in the accounting standard, for example, information about the document or entity reporting. Agencies can create their own taxonomy which can be used in conjunction with an accounting standards setter taxonomy; this was explained above in the section covering the re-use of existing taxonomies where possible.

### ***Step 2. Create Agency-specific taxonomies following the single semantic data model structure where voluntary standards organizations are not available.***

For those situations where a data collection is not based on a third-party accounting standard, Agencies can establish an inventory of data they collect by building taxonomies for each data collection that adhere to the single semantic data model as shown.



The development of a taxonomy (schema) starts with identifying all the information that may need to be reported by any reporting entity, and then codifying the characteristics of each reported fact including data type, units of measure if appropriate, period type, definition, and label. The second step is in determining the relationships between the facts reported, such as mathematical or presentation relationships. Identifying all the characteristics that are needed to unequivocally understand a fact simply requires following what is needed for the semantic data model.

### ***Step 3. Compare inventories across Agencies to identify overlap.***

Once data collections are inventoried in structured taxonomies, it is easy to query across all taxonomies to identify areas where concepts, such as Name or Address, can be consolidated to improve efficiency. This process can take place over time, as the various taxonomies mature, and need not delay the implementation of individual structured data collections. One important benefit of data standards following the single semantic data model structure is that they are designed to change with time, without “breaking” previously created data sets, and maintaining the ability to analyze data over time.

Once taxonomies have been developed in line with the semantic data model structure established, the Agencies should be confident in rolling out their own initial programs.

### ***Step 4. Maintain and develop taxonomies.***

Because reporting needs (including accounting standards) change over time, there is an expectation that whichever organization is responsible for the standard itself should be involved in revising the taxonomy to ensure that it remains current. Agencies and standard-setting

organizations will need to coordinate a process for ongoing development, and a governance structure to support it. The SEC and FASB are a good model for this approach. They coordinate efforts each year to develop a new taxonomy release (FASB) and review, approve, and publish the release (SEC). 6,000 companies roll over to a new release of the FASB GAAP Taxonomy each year.

***Step 5. FSOC agencies should continue the dialogue to ensure economies of scale.***

The Agencies should maintain an ongoing dialogue to share ideas, tools, and even validation rules. Accounting and other calculation rules can often be shared across Agencies because they are all collecting data based on a single semantic data model structure. This approach will generate efficiencies and economies of scale that should increase as Agencies, reporting entities and other data users become more proficient at working with the data, and as taxonomies mature.

## **Conclusion**

We appreciate the opportunity to provide our input, which was developed through a series of in-depth discussions with members of XBRL US. The worst outcome of the FDTA would be to expend significant taxpayer resources implementing a program that falls short of expectations and does not reap the significant benefits that a robust program can attain. With the exception of the adoption of the LEI, **the proposed rule in its current form maintains the status quo.**

With the recommendations we made, we strongly believe that the FDTA will be highly successful at meeting expectations, delivering greater transparency and accountability, and reducing costs across the federal government.

Thank you again for the opportunity to comment. XBRL US and our members urge the Agencies to focus efforts on an effective implementation. We are here to support your efforts and offer our expertise wherever we may be of assistance. Please contact me if you have any questions or would like to discuss our comments further. I can be reached at (917) 582-6159 or [Campbell.Pryde@xbrl.us](mailto:Campbell.Pryde@xbrl.us).

Sincerely,

A handwritten signature in black ink, appearing to read 'Campbell Pryde', is written over a light blue background.

Campbell Pryde, President, and CEO, XBRL US



October 21, 2024

1345 Avenue of the Americas  
27<sup>th</sup> Floor  
New York, NY 10105  
Phone: (202) 448-1985  
Fax: (866) 516-6923

Chair Gary Gensler  
Commissioner Caroline A. Crenshaw  
Commissioner Hester Pierce  
Commissioner Jamie Lizarraga  
Commissioner Mark T. Uyeda  
Securities and Exchange Commission  
100 F Street NE  
Washington DC 20549-1090

RE: Proposed Rule: Financial Data Transparency Act Joint Data Standards  
SEC File No. S7-2024-05

Dear Chair Gensler and Commissioners:

This letter is prepared in conjunction with our comment letter submitted in response to the Financial Data Transparency Act Joint Data Standards, SEC File No. S7-2024-05. It addresses specific questions raised by SEC Commissioner Pierce in the statement, Data Beta: Statement on Financial Data Transparency Act Joint Data Standards Proposal<sup>1</sup>, dated August 2, 2024. We appreciate the opportunity to respond to the Commissioner's questions which raise many important issues that are pertinent to implementation of the FDTA.

## Costs of FDTA implementation

**Question: What are the total direct and indirect costs of adopting the contemplated data standards?**

### *Cost to Agencies implementing a standards program*

It is difficult to gauge the costs of FDTA implementation for regulators given the great variability in scope of the various data collections. We can, however, point to an example of a recent regulatory implementation that may provide guidance on costs that Agencies may incur. The Federal Energy Regulatory Commission's (FERC) data modernization project<sup>2</sup> was conducted in 2021. It included project management, development of an XBRL taxonomy and (business) validation rules, training for internal staff, upgrades to FERC's electronic filing system, data consumption tools and APIs, conversion of 10 years of historical data from Visual FoxPro files,

---

<sup>1</sup> Data Beta: Statement on Financial Data Transparency Act Joint Data Standards Proposal:  
<https://www.sec.gov/newsroom/speeches-statements/peirce-statement-financial-data-transparency-act-080224>

<sup>2</sup> See USASPENDING.gov, Award Profile Contract Summary, Department of Energy:  
[https://www.usaspending.gov/award/CONT\\_AWD\\_89603019P0018\\_8960\\_-NONE\\_-NONE-](https://www.usaspending.gov/award/CONT_AWD_89603019P0018_8960_-NONE_-NONE-)



and development, testing and implementation of a deployment plan, The program has cost \$6 million to date, and is estimated to total \$7.4 million spent over six years when complete. The XBRL taxonomy for FERC reporting has ten forms with 303 individual schedules.

The FERC noted in a recent webinar<sup>3</sup>, *“We delivered this project on time and on budget... This was a glowing success for FERC because it didn’t take more time and wasn’t much more expensive than what was considered in the original spend.”*

We estimate the cost of the program for the collecting Agency per discrete fact (datapoint) collected to date, is approximately 10.35c since the program began. Maintaining the program going forward is substantially less than the initial build, so the cost per datapoint collected will decline over time.

**Question: Do these costs vary based on factors such as firm size or type of regulatory filing?**

### *Cost to reporting entities*

The cost and resources needed to prepare data in machine-readable format for an individual reporting entity, which could be a government entity, a credit union, a public company, or any entity reporting to one of the Agencies, will vary depending on a number of factors:

- The complexity and amount of data needed to be reported. For example, financial statements along with footnotes to the financials will be more costly to prepare than a simpler forms-based submission.
- The type and structure of the reporting entity, and the level of reporting required. For example, governments have multiple reporting units and often report financials representing government-wide statements, proprietary funds statements and governmental fund level statements. Furthermore, within a single set of financial statements, a government may be required to report data for separate funds and component units.
- The size of the reporting entity. Large entities with more robust budgets and staff may need additional features on tools they use to allow more than one individual to access, review features, etc.
- The number of times a reporting entity has prepared data in standardized format. The first-time preparing data in machine-readable, standardized data format will be the most time-consuming and labor-intensive with the need to select the appropriate XBRL concept (tag). Subsequent reporting periods will be significantly easier, because of the experience acquired, and because many reporting applications give entities the ability to view their previous period submission of data and update prior year numbers and narratives rather than begin from scratch. While entities do not report the same information each period, this “template” approach is common because usually entities report using the same XBRL concepts each period and simply need to update the facts.

---

<sup>3</sup> See FERC: How Regulators Use Technology to Improve Efficiency: <https://xbrl.us/events/240911/>

- Entities that report on an annual basis may face more challenges than those reporting more frequently because they only perform the reporting task once a year, compared to other entities that may report every quarter. A higher frequency of reporting reinforces learning with less chance of institutional learning loss a year later.
- Many government units today prepare financials using older systems that may be in need of modernization. There may also be variability between government reporting systems of general-purpose governments compared to special districts which are generally smaller with fewer resources.
- While every state government follows US GAAP, some states do not require their local government units to report under US GAAP. For example, local municipalities in New Jersey and Kentucky adhere to state-specific accounting standards. The FDTA does not require reporting entities to change what they report. Those non-GAAP reporting entities may face unique challenges which should be taken into consideration.

Keeping these factors in mind, cost information is available from other reporting domains that may provide insights into the ultimate cost for governments and other entities that will be tasked with reporting in the FDTA roll-out. Three examples are described below.

**Example 1. Small filers preparing forms-based data.**

For small filers submitting forms-based data to the FERC, the cost to prepare their filing such as Form 6Q below, was as low as \$1000 for four filings per year. Forms-based data does not allow the inclusion of footnotes or custom line items.

Name of Respondent IMTT-Pipeline, LLC		This Report Is: (1) <input checked="" type="checkbox"/> An Original (2) <input type="checkbox"/> A Resubmission	Date of Report 06/30/2024	Year/Period of Report End of: 2024/ Q2
<b>Comparative Balance Sheet Statement</b>				
1. For instructions covering this schedule, see the text and instructions pertaining to Balance Sheet Accounts in the USofA. The entries in this balance sheet should be consistent with those in the supporting schedules on the pages indicated. 2. On line 30, include depreciation applicable to investment in system property.				
Line No.	Item (a)	Reference Page No. for Annual (b)	Current Year End of Quarter/Year Balance (in dollars) (c)	Prior Year End Balance 12/31 (in dollars) (d)
	<b>CURRENT ASSETS</b>			
1	Cash (10)			
2	Special Deposits (10-5)			
3	Temporary Investments (11)			
4	Notes Receivable (12)			
5	Receivables from Affiliated Companies (13)	200		
6	Accounts Receivable (14)		207,199.48	(327,932)
7	Accumulated Provision For Uncollectible Accounts (14-5)			
8	Interest and Dividends Receivable (15)			
9	Oil Inventory (16)		104,172.38	184,944
10	Material and Supplies (17)			

Example 2. Small public companies prepare financials and footnotes to the financials.

An AICPA study<sup>4</sup> estimated the average price of XBRL preparation for small US companies in 2017 was \$5,476<sup>5</sup> per year to prepare four financial statement filings. Note that this figure does not include consulting charges which many filers may need early on in the process; some may no longer need consulting after the first few filings. Large reporting entities are likely to have higher costs than small entities due to the complexity of their financials and of additional features they may opt for in their reporting and disclosure management tools, for example, the ability for more than one individual to access the filings for review and editing which may involve permissioning features.

Obtaining more current XBRL-only pricing is difficult today because of the widespread use of disclosure management tools which incorporate XBRL preparation along with many other reporting management features; the XBRL tagging portion of the tool cannot be separated out.

<sup>4</sup> AICPA, XBRL costs for small reporting companies have declined 45% since 2014:

<https://us.aicpa.org/content/dam/aicpa/interestareas/frc/accountingfinancialreporting/xbrl/downloadabledocuments/xbrl-costs-for-small-companies.pdf>

<sup>5</sup> Note that this figure is not adjusted for inflation and covers the cost of the tool alone. It does not include consulting that may be required.

The practice of “bolt-on” tools, which help a company convert their corporate reports into Inline XBRL format without changing any other aspect of their reporting arrangements or workflow, was once more common but is increasingly being replaced by disclosure management tools. Trends in the use of disclosure management tools are addressed later in this letter.

### Example 3. Small businesses report tax information.

Her Majesty's Revenues & Customs (HMRC) Companies House<sup>6</sup> in the United Kingdom is one regulator that requires over 4.5 million companies<sup>7</sup> to report tax data in structured (XBRL) format today. Companies House reporting entities include doctor's offices, drug stores, restaurants, and other small businesses, which are likely to be as resource-constrained as small government entities in the U.S. The commercial market of accounting applications already used by these small businesses, has addressed the issue in large part by incorporating XBRL preparation and export into their applications.

### How have standardized data costs changed over time? Will these trends continue?

The SEC program for public companies provides insights into costs and resource requirements for other types of reporting entities such as governments, credit unions, or banks. As noted earlier, costs tend to be higher the first time an issuer prepares data in structured format given the learning curve, and because the second time preparing data in structured format, the issuer can build on the work completed during the first filing period. For example, the appropriate “tags” have already been identified, and software usually allows reporting entities to view and revise the previous set of data reported as the XBRL “tags” used for one period are often the same the following quarter or year.

Pricing has also declined for structured data (XBRL) reporting tools and services. The AICPA study<sup>8</sup> referenced earlier was conducted in 2014 and again in 2017. 2017 estimates were 45% lower than estimates in 2014.

Costs are likely to continue to decline because of the expansion in regulatory reporting requirements calling for data standardization. The FERC, for example, began a new data standards (XBRL) program for utilities in 2021; and they are soon to introduce another XBRL program calling for XBRL prepared in CSV format (XBRL-CSV). Products that prepare data in Inline XBRL can also prepare data in XBRL-CSV, or XBRL-JSON, or XBRL-XML.

Separately, European publicly listed companies today report financial data in XBRL format, and starting in 2026, both publicly listed and private companies in Europe will be reporting climate data to European Union (EU) authorities in structured format. This program is phasing in between

---

<sup>6</sup> UK HM Revenue & Customs, XBRL guide for businesses: <https://www.gov.uk/government/publications/xbrl-guide-for-uk-businesses/xbrl-guide-for-uk-businesses#introduction>

<sup>7</sup> UK Companies House Official Statistics, Incorporated companies in the UK April to June 2024: <https://www.gov.uk/government/statistics/incorporated-companies-in-the-uk-april-to-june-2024/incorporated-companies-in-the-uk-april-to-june-2024>

<sup>8</sup> AICPA, XBRL costs for small reporting companies have declined 45% since 2014: <https://us.aicpa.org/content/dam/aicpa/interestareas/frc/accountingfinancialreporting/xbrl/downloadabledocuments/xbrl-costs-for-small-companies.pdf>

2026 to 2029 with the largest public companies in the first phase. More than 20 countries that follow the International Financial Reporting Standard (IFRS) are working towards adoption of digital (XBRL) climate reporting as well, using the International Sustainability Standards Board (ISSB) Taxonomy. The first of these climate reporting mandates is expected in 2026 as well.

The increasing demand for XBRL data, and thus, XBRL tagging tools, has already initiated an expansion of the market for tools, which in turn leads to greater competition. This drives down pricing, and encourages the development of more streamlined, efficient applications. The economies of scales enabled by standardization is one of the primary reasons they are so effective.

When public companies began XBRL preparation in 2009, many relied on outside consultants. Today, most have transitioned to disclosure management applications which include XBRL preparation; these tools have also introduced quality, control, and efficiency enhancements that make the entire reporting and preparation process more effective. According to a PwC report<sup>9</sup> on integrated reporting, the increasing use of disclosure management tools, sometimes called the “last mile” of reporting, has been driven by expanding financial and non-financial (for example climate) reporting requirements, global capital financing, the expanding use of Artificial Intelligence, and the partial or total migration that many organizations have made to cloud technologies.

The disclosure management market is estimated at \$1.2 billion in 2024 and is expected to grow at a CAGR of 16.5% by 2029, as noted in the report, *Disclosure Management Market Size & Share Analysis - Growth Trends & Forecasts (2024-2029)*<sup>10</sup>. The report goes on to note, *“The North American disclosure management market is expected to grow significantly, owing to technologically advanced countries like the United States and Canada. Companies are dealing with increasing complexity in the reporting process characterized by different file formats, causing the demand to rise. The quicker report preparation and financial statement disclosure, report comparison, tracking changes made by any user, and complete restoration of earlier versions provide greater control to the users, propelling demand in the market.”*

This viewpoint is echoed in a 2023 Journal of Accountancy article<sup>11</sup> which notes that *“...investing in technologies that automate core processes and streamline user experience will be paramount to building - and retaining - a skilled and agile finance team.”*

Corporate entities are leading the way in embracing disclosure management tools and governments are beginning to follow suit as they recognize how they can benefit too. Workforce shortages in public sector accounting and finance are also driving the move towards greater

---

<sup>9</sup> PwC, The need for integrated reporting during times of change: <https://www.pwc.com/gx/en/services/audit-assurance/capital-market/digitising-finance/disclosure-management.html>

<sup>10</sup> Mordor Intelligence, DISCLOSURE MANAGEMENT MARKET SIZE & SHARE ANALYSIS-GROWTH TRENDS & FORECASTS (2024-2029): <https://www.mordorintelligence.com/industry-reports/disclosure-management-market>

<sup>11</sup> Journal of Accountancy, “3 trends that will reshape accounting and finance in 2023”, March 15, 2023: <https://www.journalofaccountancy.com/news/2023/mar/3-trends-reshape-accounting-finance-2023.html>



automation through technology. As noted in a recent Governing Magazine article<sup>12</sup>, *“Departments are struggling to recruit and retain the next generation of workers, leading to lags in reporting, big backlogs for processing local taxes and licenses and a lack of auditors. The workers who are still around are getting older.”* This sentiment was echoed in a Government Finance Officers Association (GFOA) report<sup>13</sup> from 2022, which noted, *“Demand for state and local public finance officers is outstripping the current supply of workers in the sector.”* Greater use of technology and automation to conduct more effective research can reduce the hiring challenge because it improves productivity on the job, but even more importantly it can make public sector finance jobs more attractive to younger, tech-savvy recruits.

The advent and rapid adoption of disclosure management tools has been facilitated by the existence of standardized data. The benefits of these applications extend well beyond XBRL preparation requirements. Disclosure management tools allow reporting entities to review prior period reported facts and narrative disclosures, so that they can revise facts with updated information efficiently, not to encourage boilerplate submissions.

Smaller reporting entities, however, may not have the resources or even the need for disclosure management solutions. Small governmental units may prepare their financials in a spreadsheet that is then converted to PDF, or in many cases, rely on their audit firm or a consulting firm to prepare their annual financial report or annual comprehensive financial report. This is not allowed for publicly traded companies. Different pricing models that support a large market of small, resource-constrained entities already exist in other reporting domains and provide a helpful model that can be used for FDTA programs where small entities must be considered. As mentioned earlier, HMRC Companies House<sup>14</sup> is one regulator that requires over 4.5 million companies<sup>15</sup> to report tax data in structured (XBRL) format today. The accounting software used by these small businesses has XBRL export features to allow them to comply efficiently and cost-effectively.

**Question: How much of the FDTA compliance burden is likely to stem from the one-time cost of setting up new data systems as opposed to ongoing compliance costs?**

Reporting entities preparing their data in structured, standardized format for the first time are likely to have a learning curve as they work to understand the taxonomy and how to identify the appropriate concepts to match the data they are reporting. Documentation like web-based tools for viewing the hierarchical structure of taxonomy concepts aid in this process (i.e.

---

<sup>12</sup> See Governing Magazine, July 3, 2024, The Biggest Challenge in Public Finance May Not Be Money: <https://www.governing.com/workforce/the-biggest-challenge-in-public-finance-may-not-be-money#:~:text=Workforce%20shortages%20are%20affecting%20many,few%20candidates%20replacing%20aging%20employee s.&text=In%20Brief%3A,still%20below%20pre%20pandemic%20numbers>.

<sup>13</sup> See GFOA, September 2022, Meeting Demand for State and Local Public Finance Jobs: [https://gfoa.org.cdn.prismic.io/gfoaorg/5c3e1746-08da-43fa-af63-1445e97db05b\\_GFOA-Lightcast-Report.pdf](https://gfoa.org.cdn.prismic.io/gfoaorg/5c3e1746-08da-43fa-af63-1445e97db05b_GFOA-Lightcast-Report.pdf)

<sup>14</sup> UK HM Revenue & Customs, XBRL guide for businesses: <https://www.gov.uk/government/publications/xbrl-guide-for-uk-businesses/xbrl-guide-for-uk-businesses#introduction>

<sup>15</sup> UK Companies House Official Statistics, Incorporated companies in the UK April to June 2024: <https://www.gov.uk/government/statistics/incorporated-companies-in-the-uk-april-to-june-2024/incorporated-companies-in-the-uk-april-to-june-2024>

<https://xbrlview.fasb.org>) and often, report preparation software also includes this kind of functionality. Subsequent reporting efforts are significantly easier, particularly when entities are tasked with reporting the same information each year or quarter, for example, governments submitting an Annual Comprehensive Financial Report, or banks preparing quarterly financial statement filings.

Widely used disclosure management tools give preparers the means to re-use an XBRL preparation “template” that reflects their prior year or quarter so they can update an existing template with new numbers and new narratives for the next reporting period. Because the data structure is standardized, subsequent report preparation is streamlined, as the structure of the report and the data from the prior period is often rolled-forward. This represents an improvement in terms of time savings compared to the multi-stage process that occurs today, where a report may be prepared in Excel or Word, combined into a single document, and then converted to PDF.

Structured data also enables automated validations for accounting rules (checking that totals and subtotals are correct, that a fact reported on one statement matches the same fact reported on another statement, etc.), reasonableness, completeness, and other business rules that ensure higher data integrity. This approach is followed today by public companies and public utilities through the SEC and FERC programs, respectively. Validation rules can improve the efficiency of report preparation as they can automatically point out inconsistencies automatically.

**Question: Will certain types of entities, such as municipal issuers, bear disproportionate FDTA-related costs? If so, what can we do to reduce those costs?**

Initially, reporting entities like municipal issuers, as the first “link on the reporting chain” will bear greater costs of data standards implementation although this will decline over time as it has for corporate issuers. The immediate beneficiaries will be regulators and other data users.

Regulators and the market can help issuers by considering different cost models as discussed earlier. The federal government or state regulators could negotiate the bulk purchase of low-cost software tools for the smallest issuers. This approach could provide an easier on-ramp for small entities; however, we encourage regulators that take this approach to ensure reporting entities are able to choose from a variety of options to maintain the competitiveness of the market.

FASB accounting support fees<sup>16</sup> are assessed on and collected from issuers of publicly traded securities, as those issuers are defined in the Sarbanes-Oxley Act and are allocated based on the average market capitalization of each issuer. These funds support the accounting standard setting process as well as the XBRL taxonomy development process. The FASB Project Staff, which is responsible for setting accounting standards, works closely with the FASB Taxonomy Staff developing accounting standards, modeling them into the FASB GAAP Taxonomy and publishing them together during a public exposure period (read more about this approach<sup>17</sup>). This process

---

<sup>16</sup> FASB Accounting Support Fees: <https://accountingfoundation.org/page/pageContent?pagelId=/about-us/faf-fasb-gasb-how-we-are-funded/accounting-support-fees.html>

<sup>17</sup> Harmonizing accounting and data standards: <https://xbrl.us/harmonizing-accounting-data-standards/>

has been in place since 2018 and it allows the Taxonomy Staff to inform the Project Staff about the impact that improvements to accounting standards may have on the Taxonomy, on constituents and on the resulting data. An accounting support fee<sup>18</sup> is also assessed by FINRA to fund the annual budget of the GASB. FDTA compliance could be supported by an addition to this fee.

During a recent event conducted by XBRL US in July 2024, we solicited input from market participants about the FDTA roll-out during brainstorming sessions<sup>19</sup>. Participants felt that there was an opportunity for different cost models to be developed, for example it was suggested that the federal government should consider creating a platform that helps the smallest local governments comply easily and at low cost. The IRS' Direct File<sup>20</sup> pilot - albeit for individual taxpayers - illustrates a means to allow certain filers access to a no-cost solution. While this may be challenging for complex reports like financial statements, it shows a creative solution offered by the government to keep costs low for certain kinds of reporting. Phase-ins could be considered as well based on size of entity or type of information reported.

Providing a “no-frills” solution at zero cost is a good step, however we urge regulators to encourage commercial applications to be made available as well. Constraining reporting entities to using a single application eliminates the benefits of a competitive marketplace. Access to multiple options for reporting ensures the lowest possible cost, and robust tools for reporting entities.

## **Benefits of FDTA implementation.**

**Question: What are the potential benefits of FDTA implementation?**

### ***Access to more granular, automated data for all reporting entities***

With proper implementation of a single semantic data model structure, the FDTA will result in data that is more timely and easier to process because it can be consumed automatically and understood by machines. Investors and data analytics providers have stated that they prefer public company data in XBRL because it offers certain benefits not found in other datasets. These benefits are supported by quotes from several data and analytics providers that were recorded during a video interview<sup>21</sup>. While the quotes refer to SEC reported data from public companies, the same benefits can be gained by municipal analysts and investors if government data were available in the same format:

---

<sup>18</sup> FINRA Regulatory Notice 24-07: 2024 GASB Accounting Support Fee to Fund the GASB: <https://www.finra.org/rules-guidance/notices/24-07#:~:text=The%20GASB%20Accounting%20Support%20Fee%20is%20assessed%20on%20a%20quarterly,Fee%20from%20its%20member%20firms.>

<sup>19</sup> Blog, Muni Bond Markets Weigh in at GovFin 2024: <https://xbrl.us/govfin2024-muni-market-attendee/>

<sup>20</sup> See IRS Direct File: <https://www.irs.gov/newsroom/irs-makes-direct-file-a-permanent-option-to-file-federal-tax-returns-expanded-access-for-more-taxpayers-planned-for-the-2025-filing-season>

<sup>21</sup> Video: Better Data for Analysts and Investors: <https://xbrl.us/news/analyst-video/>

- Improves access to small issuer data. This can be especially beneficial to small government municipal bond issuers that have little visibility. Fund managers are generally compelled to focus on large issuers to meet client demand, often to the detriment of smaller issues. With data from small and large entities available in the same machine-readable format and at the same level of granularity, there is a greater opportunity for small issuers to gain visibility and investors:  
*"... [with XBRL] there's no difference in the availability of data between large and small companies." Pranav Ghai, CEO, Calcbench*
- Allows access to more detailed, granular data:  
*"[accessing data in the footnotes to the financials] is not that easy to do unless you have access to structured data, in this case XBRL...with data available in the XBRL format, we can extract data from the footnotes in seconds. From one company, from thousands of companies." Pranav Ghai, CEO, Calcbench*
- Increases ease of data processing. Standardized, structured data enables automation which ultimately results in data being made available to governments, regulators, researchers, and citizens less expensively with greater granularity and potentially higher data integrity:  
*"Extracting data from an HTML document takes at least 20 minutes, from a good quality PDF, takes around 30 minutes, from an image around 50 minutes. Data pulled from an XBRL file, though, can be extracted in 1 to 2 seconds... let us focus on better analytics rather than scraping data from documents." Adrien Cloutier, Global Director of Equity Data, Morningstar*

Academic research has shown that as-reported corporate data in XBRL is more timely, granular, and authoritative, and has greater predictive qualities than data that must be manipulated, standardized, and normalized by data providers. A study<sup>22</sup> conducted by Penn State University found that XBRL data is better at predicting future stock returns than commercially created data that is manipulated to become interoperable.

Today, government financial data is only available by extracting data from PDFs or by purchasing costly data licenses from commercial entities (which have performed the labor-intensive data scraping, manual vetting, and normalization necessary today to ready the data for use). When a single semantic data model structure is implemented, businesses, governments, and researchers will have access to granular, inexpensive data as well.

When this data is freely available from regulators in structured, standardized form, anyone can extract and use the data. For example, a US-based insurance company transitioned away from manual data entry of corporate financial data they use for benchmarking and analysis, to directly accessing XBRL data from the SEC which the company now automatically extracts into its internal

---

<sup>22</sup> Are XBRL data better at predicting future stock returns? Smeal College of Business, Penn State University:  
<https://xbml.us/xbml-better-at-predicting/>

data analytics platform. They found significant benefits from access to more granular, timely data, and from redirecting employee time to higher value work. They estimated a 150% ROI in year one of the implementation<sup>23</sup>.

#### ***Better data integrity and processing efficiency gains.***

The Federal Deposit Insurance Corporation (FDIC) began collecting data from banks using a single semantic data model (XBRL) in 2005. Their implementation resulted in immediate efficiency gains for members of the Federal Financial Institutions Examination Council (FFIEC)<sup>24</sup> including<sup>25</sup>:

- 95% of data received met FFIEC validation requirements.
- 100% of data received met FFIEC mathematical validation requirements versus 70% in the legacy system.
- Data was publicly available immediately after calendar quarter end, versus weeks later in the prior legacy system.
- Staff productivity in handling bank data increased 10-33%.
- Data could be distributed to end users at agencies within one hour versus within several days in the legacy system.

#### ***Economies of scale for reduced cost, more competition, better tools for all***

When the semantic data model structure is used, software applications become more commoditized and interchangeable. Software companies do not have to build custom products and processes for each data collection. One tool can support multiple types of reporting entities and reporting requirements. The software company's cost of tool development and maintenance can be shared across many reporting entities and the competitive nature of the market means that those savings are passed down to reporting entities.

Data collected can be shared, inventoried, commingled, and compared across states, giving governments significantly greater insights into how to manage budgets, set policies, share ideas, and identify which programs are more effective than others. Taxpayers can more easily learn how their governments are spending money, and which programs are performing well. They can hold governments accountable for what they spend.

Lessons can be learned from the SEC's public company implementation of XBRL reporting. When corporate data became available in XBRL format, many new data and analytics companies came on the market because the structured data was freely available and could be automatically extracted and used without the need for manual data entry and review. This opening of the market lowered the cost of data to all data consumers.

The proposal as written however, gives Agencies the flexibility to create custom schemas which will result in data sets that are not interoperable. Data sets that are not interoperable require significant mapping and manipulation to enable comparison, sharing, or searching in the same database environment. The cost of data manipulation adds to the cost of analysis, and often

---

<sup>23</sup>Custom Data Collection with Standardized Data: Liberty Mutual Surety Case Study: <https://xbrl.us/research/data-collection/>

<sup>24</sup> See FFIEC: <https://www.ffiec.gov/about.htm>

<sup>25</sup> See FFIEC, Improved Business Process Through XBRL: A Use Case for Business Reporting: <https://xbrl.us/wp-content/uploads/2007/12/20060202FFIECWhitePaper.pdf>



means that data from large entities is extracted and used, while small entities are ignored because it's just too costly to extract and manipulate the entire universe of securities. Our comment letter on the rule proposal outlines our objections to this approach and recommendations on how the FDTA can be successfully implemented.

**Question: How could the Commission maximize the utility of financial regulatory information filed in compliance with the FDTA-mandated data standards? For example, should the Commission work to reduce error rates in structured data filings?**

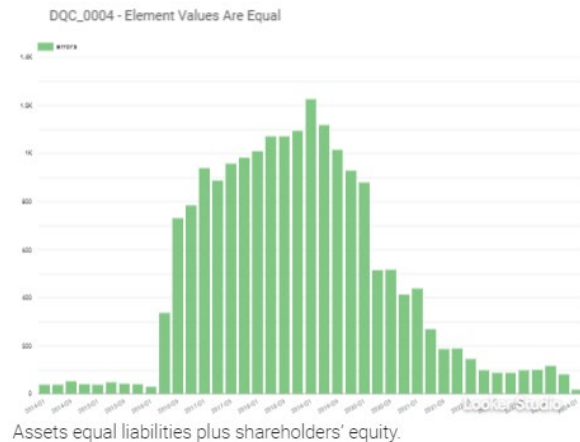
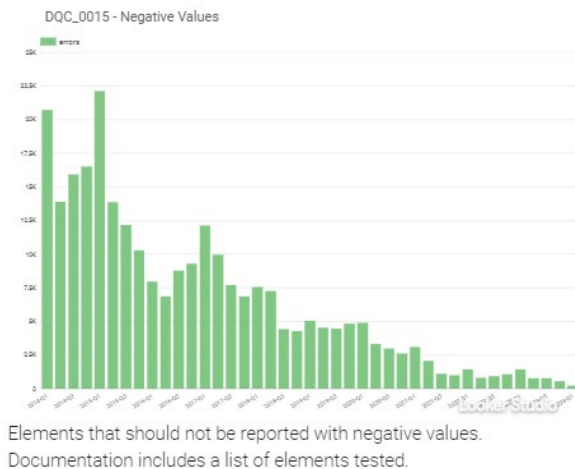
Standardized, structured data produced through a rigorous implementation of the FDTA will lead to much easier review and vetting. Automated validation rules can be created to identify inconsistencies which can be corrected prior to submitting reports to the regulator. The nature of structured data enables the creation of very detailed rules that can instantly check for errors and inconsistencies such as values that should be reported with a particular sign (positive or negative), facts that are required to be reported and missing, facts that should or should not be reported when another fact is reported, as well as reasonableness checks and assertions that ensure that accounting standard rules are followed correctly. Automated rules can be used in conjunction with attestation checklists already in use.

Precedent for this process exists today for public company reporting to the SEC and utilities reporting to the FERC. Public companies have had access to freely available data quality checking rules since 2015; these rules are now incorporated into the US GAAP Taxonomy, leveraged by most reporting applications used by public companies, and triggered in the SEC EDGAR System when filings are submitted.

These rules have improved the quality of data reported as issuers use them to correct problems during the report creation and review process. Tracking of these errors over time shows the steep decline in errors since rules were introduced and implemented.<sup>26</sup> The graphs below show the decline in errors since the rules were first introduced for facts that were reported as negatives but should have been positives, on the left; and for facts reported that did not meet the accounting calculation that Assets must always equal Liabilities plus Shareholders' Equity.

---

<sup>26</sup> Aggregated Real-time Filing Errors: <https://xbrl.us/data-quality/filing-results/dqc-results/>



These rules were created by a nonprofit industry-driven consortium called the Data Quality Committee of the Center for Data Quality<sup>27</sup> and are used today by most SEC reporting entities that use the US GAAP and the IFRS XBRL Taxonomies. Similar error reductions are found for many other rule types that the Data Quality Committee has created and published.

They are developed and maintained on a small budget and made freely available to all issuers. This approach could be adopted for government reporting as well, as an industry-driven initiative that coordinates its efforts with the tools market, the standard setter, and the reporting community.

Government financial statements are complex and can follow different accounting principles; and reported facts may appear on more than one statement. For example, many items on the Statement of Net Position (for government-wide financials) should match certain items on the Proprietary Funds Statement of Net Position, unless there is a required adjustment. Data in unstructured format, like government financial statement data reported in PDF files today, can only be validated through manual checking. Vetting, if any is performed, is likely only as a “spot check” and cannot be scaled up to cover many filings. The availability of this data in structured format would allow **every** financial statement to be automatically checked for a base level of data accuracy. Rules can be established that alert reporting entities of any discrepancies so they can be corrected before sending on to regulators, citizens, and investors.

<sup>27</sup> See Data Quality Committee: <https://xbrl.us/data-quality/center/committee/>

**Question: In light of the Commission's broad discretion in implementing the second phase of FDTA rulemaking, are there areas where full application of these proposed data standards would be particularly helpful or unhelpful?**

The municipal bond issuance market will gain from data standards implementation. Bond issuers submit a range of information to the Municipal Securities Rulemaking Board (MSRB) Electronic Municipal Markets Analysis (EMMA) system. During the focus group-type sessions mentioned earlier that were conducted at the XBRL US GovFin 2024 conference, participants reported that financial statements, official statements, material event notices, broker dealer disclosures and portions of the audit report were of great interest to data users and would be more valuable and useable in structured format.

Standardizing this data, and use of the LEI in conjunction with data standardization can also improve the ability to search on the EMMA system. Today, the ability to link the obligor, the issuer, and the security is extremely challenging because of the paper-based structure of the content submitted to EMMA. Use of the LEI combined with the linking characteristics of the XBRL standard could resolve that issue and vastly increase the efficiency of using the EMMA system. This approach is explained further in the XBRL US blog and detailed paper, Identifying the Obligor for Municipal Securities<sup>28</sup>.

The single semantic data model structure within which the Agencies develop and share schemas for collecting information will result in long-term benefits to filers and regulators by reducing redundancy in reporting and data governance. Beyond the time and material costs of transition, the quality of data will improve while its cost to produce and manage collection will decline. Benefits will include:

- Data produced will be machine-readable and machine-understandable, eliminating the need for manual data entry and vetting.
- Data will be interoperable and shareable and can be inventoried and maintained together in a database.
- The opportunity for automated checking, which can review all data reported and can be performed consistently and at a low cost.
- Economies of scale will reduce cost for regulators, reporting entities and data users.
- Flexibility for Agencies to transition to new data transmission formats in future, without being locked into particular technologies
- Enables easier updating of reporting requirements by the accounting standard setter or the Agency

**Question: What could the benefits of interoperability be in the intermediate and longer term? What level of interoperability is necessary to achieve these benefits?**

Interoperable data based on a single semantic data model structure can be readily commingled, shared, and inventoried; and can be managed using the same applications to prepare, collect,

---

<sup>28</sup> See Identifying the Obligor for Municipal Securities: <https://xbrl.us/identify-municipal-obligor/>

extract, and analyze. These attributes of interoperable data can reduce data management costs for the federal government and for every other data user; and will facilitate access to more consistent, granular, and understandable data. XBRL US, for example, maintains a database that is updated every 15 minutes with financial data from public utilities, public companies and companies reporting to the European Union<sup>29</sup>. This is possible with a small technical staff because of the structured nature of the data reported.

Reported data is either interoperable or it is not interoperable. The attempt to map datasets to each other is costly and labor-intensive, and the mapping process is likely to introduce errors and require manual review to ensure data integrity. Mapping must be performed by every data user and data collector, thus exponentially increasing the level of work required. Furthermore, separate mapping exercises may result in inconsistencies across datasets and no single source of “truth.”

**What are the largest hurdles to interoperability of financial regulatory data across financial regulators? How should we address those hurdles?**

The biggest hurdle to interoperable data is allowing data to be reported using custom designed schema; or in not requiring a schema for data at all, but in permitting data to be conveyed in data transmission formats such as CSV or PDF/A, as is currently proposed.

Collecting data using multiple custom schemas is problematic as the schemas will need to be mapped to achieve some measure of interoperability. This is explained in greater detail in the rule proposal comment letter. Reporting data in data transmission formats (like CSV) without an associated schema or taxonomy may produce machine-readable *documents*, but not machine-readable *data* that can be unambiguously understood from computer to computer. These hurdles can be addressed if data is collected using a single semantic data model structure which will result in data that is interoperable and a process that is efficient.

**Question: How could we achieve the benefits of interoperability without imposing unnecessary costs on reporting firms, particularly smaller ones?**

As noted earlier, different cost models should be considered to assist small governments or other entities that do not have the resources or staff that large organizations do. State regulators could engage multiple software providers for a bulk purchase of licenses at a low cost that could be used by smaller government entities or develop a web-based filing process with limited functionality for these entities, as the IRS has done with some individual taxpayers.

We wish to emphasize however, that selecting a single provider will not reap the benefits that the FDTA program promises; regulators should encourage competition among vendors so that pricing stays low, and the quality of applications is high, and vendors have incentive to have the highest quality performances. Standardization programs at scale will encourage general ledger providers

---

<sup>29</sup> See <https://xbrl.us/home/use/filings-database>

to incorporate standardized data export modules into their tools which will minimize added cost and disruption to reporting entities.

**Question: Did the joint regulators strike the right balance when choosing the proposal's scope, in light of the statutory mandate?**

We understand that the joint regulators sought to balance mandatory requirements with flexibility for Agencies working with different data collections. That said, the joint rule as written provides too much latitude in the properties-based approach. The proposed rule allows Agencies to select from a menu of choices that will not produce data that is interoperable, and are in some cases, not data standards at all (XML, CSV, PDF/A and JSON are data formats, not data standards).

**Question: What policy issues should the Commission consider when adapting its rulebook to conform to the joint data standards, particularly given the statutory discretion afforded to the Commission (and other implementing agencies) in the second-phase rulemaking.**

The Commission should consider how it can ensure interoperability among its own datasets by adopting a single semantic data model structure. Today, as noted in the responses to the Joint Agency rule proposal, the SEC requires companies that report using the US GAAP accounting standard to do so, in three different ways. This would be a good time to transition to adoption of a single semantic data model for all companies reporting in US GAAP. Companies that are subject to Regulation A and Regulation Crowdfunding adhere to the US GAAP accounting standard, but currently report using two different custom XML schemas. Public companies report in XBRL format using the US GAAP Taxonomy maintained by the Financial Accounting Standards Board (FASB).

**Question: Does the proposal strike the right balance between principles and prescription? If not, how should the approach change?**

As noted in our response to the Joint Agencies rule, the properties-based approach offers far too much latitude to Agencies to adopt data transmission formats that will not produce interoperable data. We urge the Agencies to reconsider this approach and implement a single semantic data model structure that supports a variety of data transmission formats. This will ensure that all Agency data collections are interoperable; it will give the Agencies flexibility to choose from widely used data transmission formats; and it will provide the means to evolve the approach over time if and when new data transmission formats are developed and can be easily incorporated into Agency ongoing rulemaking.

The Agencies should also consider how other non-FSOC agencies collect and use data today from the same reporting entities. For example, the U.S. Census uses government data extensively; many other agencies collect grants data from governments as well. Ensuring that a single approach is adopted across all these agencies would enable interoperability and even greater economies of scale and reduced cost.



**Question: Would the balance the proposal strikes allow data standards to be updated in a timely manner? If not, what would work better? How often should regulators revisit the mandated standards to ensure that they remain current? Should we build a requirement to revisit the standards into the final rule?**

The property based standard proposed does not provide a mechanism to easily expand or build upon the taxonomies. The single semantic data model structure provides a modular, long-term solution with the rigid structure to generate consistent data but the flexibility to continuously update and revise standards, technologies, and reporting needs quickly and efficiently.

When there are changes in accounting standards and industry requirements, taxonomies can be revised with relative ease and minimal need for IT support. When new data transmission formats are introduced, the semantic data model structure can be adapted to transport data in the new format with relative ease.

For example, the taxonomy supporting US public companies is updated with a new release each year that reflects changes in accounting standards and industry requirements; 6,000 public companies transition to a new release with minimal effort. The FDIC Taxonomy for banks is sometimes updated on a quarterly basis, with 5,000 banks transitioning to a new release, again with minimal effort.

**Question: How, if at all, will artificial intelligence or other technologies influence the need for structured data? How should we take these potential future developments into account in implementing the FDTA?**

A recent Department of Commerce publication, AI and Open Government Data Assets Request for Information<sup>30</sup> noted, *“today’s AI systems are fundamentally limited by their reliance on extensive, unstructured data stores, which depend on the underlying data rather than an ability to reason and make judgments based on comprehension.”*

The RFI aimed to explore how to achieve better data integrity, accessibility, and quality because *“AI tools are increasingly used for data analysis and data access, so Commerce hopes to ensure that the data these tools consume is easily accessible and “machine understandable” versus just “machine readable.”* “

We agree that AI systems will be strongly supported by the availability of structured, standardized, interoperable datasets envisioned by the FDTA.

A recent blog post, The IT Leader’s Guide to Preparing Structured and Unstructured Data for Generative AI<sup>31</sup> noted, *“Imagine you’re a chef trying to cook a complex dish in a kitchen where all*

---

<sup>30</sup> See U.S. Department of Commerce, AI and Open Government Data Assets Request for Information : <https://www.govinfo.gov/content/pkg/FR-2024-04-17/pdf/2024-08168.pdf>

<sup>31</sup> See The IT Leader’s Guide to Preparing Structured and Unstructured Data for Generative AI: <https://shelf.io/blog/the-it-leaders-guide-to-preparing-structured-and-unstructured-data-for-generative-ai/>

*the ingredients are mixed up in a big pile. This chaotic kitchen is like unstructured data: everything is there, but it's all jumbled together, making it hard to find what you need, or even know what is really there in the first place. Unstructured data, like random piles of ingredients, can still be useful. It's like having a variety of foods from different recipes all over your kitchen. But, if you don't manage this unstructured data – say, by sorting out what ingredients are there and how they can be used – you can't effectively incorporate them into your dish."*

The post goes on to note that data, like ingredients, needs structure, "*Structuring data is beneficial for both the training and the application of generative AI.*" It enables efficient data processing, supports the ability to simplify and convert complex data into a format that is easier for computer programs to understand, and it helps to identify relationships between data types for trend and correlation analysis. Structured data supports scalability and reusability, allowing an AI model to train and adapt. Structuring data can be accomplished through "tagging" which is how XBRL associates' information with a fact; through cataloging, and through the use of metadata to clarify meaning.

Artificial intelligence holds enormous promise, especially if it is supported by access to fully interoperable, structured data that can result from the successful implementation of the FDTA.

With the recommendations we made in our FDTA comment letter, we strongly believe that the FDTA will be highly successful at meeting expectations, delivering greater transparency and accountability, and reducing costs across the federal government. Thank you again for the opportunity to comment. Please contact me if you have any questions or would like to discuss our comments further. I can be reached at (917) 582-6159 or [Campbell.Pryde@Xbri.us](mailto:Campbell.Pryde@Xbri.us).

Sincerely,

A handwritten signature in dark ink, appearing to read 'C. Pryde', is positioned below the word 'Sincerely,'.

Campbell Pryde, President and CEO, XBRL US

CC: Ann E. Misback, Secretary, Board of Governors of the Federal Reserve System  
Christopher Kirkpatrick, Secretary, Commodity Futures Trading Commission  
Legal Division, Docket Manager, Consumer Financial Protection Bureau  
James P. Sheesley, Assistant Executive Secretary, Federal Deposit Insurance Corp.  
Clinton Jones, General Counsel, Federal Housing Finance Agency  
Melane Conyers-Ausbrooks, Secretary of the Board, National Credit Union Admin.  
Chief Counsel's Office, Office of the Comptroller of the Currency  
Vanessa A. Countryman, Secretary, Securities and Exchange Commission  
Chief Counsel's Office, Department of the Treasury