July 1, 2021

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

James P. Sheesley  
Assistant Executive Secretary  
Attention: Comments-RIN 3064- ZA24  
Federal Deposit Insurance Corporation  
550 17th Street NW  
Washington, DC 20429

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

Comment Intake  
Bureau of Consumer Financial Protection  
1700 G Street, NW  
Washington, DC 20552

Melane Conyers-Ausbrooks  
Secretary of the Board  
National Credit Union Administration  
1775 Duke Street  
Alexandria, VA. 22314-3428.

RE: Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning (Docket No. OCC-2020-0049; OP-1743; RIN 3064- ZA24; CFPB 2021-0004; NCUA 2021-0023)

To Whom It May Concern:

Thank you for the opportunity to provide input to the Request for Information on Financial Institution Use of Artificial Intelligence (AI), including Machine Learning, jointly requested by the Department of the Treasury Office of the Comptroller of the Currency, Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, Bureau of Consumer Financial Protection, and National Credit Union Administration.
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 eXtensible Business Reporting Language (XBRL). XBRL is a free and open data standard widely used in the United States, and around the world, for reporting by public and private companies, as well as banks and government agencies.

As a data standards organization, we strongly support digitization and technology to improve efficiencies in the preparation, collection, extraction, and analysis of financial and other types of data. The XBRL financial data standard has helped regulators and businesses worldwide,¹ in 180+ regulatory programs, automate their processes to increase efficiency and timeliness of access, and improve accuracy and granularity of information needed for decision making.

Artificial intelligence is driven by machine learning, the practice of using computer algorithms to build a model based on large amounts of “training data”. Patterns identified in training data are then used to make predictions or decisions. The most critical driver behind conclusions drawn through AI platforms is the data. For an AI system to effectively extract, understand, analyze, and learn from vast quantities of data, requires access to data that is of good quality, and that is clearly and unambiguously defined. An AI system needs to review lots of examples, drawn from lots of data, and supported by an open data ecosystem with unambiguous standards for structuring and labeling that data. The quantity and quality of the data supporting an AI system is a critical factor in its usefulness.

While this Request for Information focuses on financial institution use of AI platforms, we want to point out that regulators should also consider how financial data that the regulator collects from these institutions is used. Each of the five federal agencies that is part of this Request collects financial data from multiple financial institutions. Data aggregators, analytics providers, and services offering AI platforms access the data provided by these regulators.

We urge the regulators that are part of this Request for Information to ensure that the data they collect from financial institutions is equally available to all users, of high quality, and available in structured, standardized format. Providing financial institution data in machine-readable, structured format will ensure that it is easier and less expensive to analyze; and that AI platforms that leverage it, produce better, more reliable outcomes not just for regulators but for the public as well.

¹ The XBRL standard is used in 184 regulatory implementations worldwide to support capital markets, tax authorities, business registrars, financial regulators, government oversight programs and more: https://www.xbrl.org/the-standard/why/xbrl-project-directory/.
Why AI needs standardized data

XBRL was developed to handle the complexity of financial data, and can also be used to represent data types that are often commingled with financial data, like percent, duration, power, length, energy, boolean, integer, volume, mass, and area.

In the U.S., entities\(^2\) reporting to the Securities and Exchange Commission (SEC), banking institutions reporting to the Federal Deposit Insurance Corporation (FDIC), and public utilities reporting to the Federal Energy Regulatory Commission (FERC), are required to report in machine-readable (XBRL) format. Because they are required to report using the same data standard, the same tools used to prepare, collect, extract, or analyze banking data submitted to the FDIC can be used with financial or energy-related data submitted by public utilities to the FERC, or corporate filings submitted to the SEC. Although the information reported (to the SEC, FDIC, or FERC) is quite different, the structure of the data is the same. This uniform structure ensures that the cost to preparers, regulators, data intermediaries, and data users is low because stakeholders can leverage a large, competitive marketplace of applications to prepare, collect and analyze different kinds of data.

From an AI perspective, consistently structured, machine-readable data is the fuel driving more accurate outcomes. Consider the age-old “garbage in, garbage out” concept. The predictions or decisions made by an AI platform are only as good as the data that supports the system.

If, for example, an AI platform was being “trained” by looking at various companies reporting balance sheet items, it would be critical that the platform had an unambiguous understanding of every fact reported in every balance sheet, including for example, the value 3,405 that we see on the table below. Consider the dimensional metadata associated with that single fact. It represents:

- Cash and cash equivalents
- Millions of US dollars
- For the period September 20, 2020
- Hilton Worldwide Holdings

\(^2\) SEC reporting entities required to report financials in XBRL format include public companies, investment management companies, credit rating agencies, business development companies, and variable annuity and variable life insurance companies.
Financial data, by its nature, is complex and can have even more associated metadata than this fact reported by Hilton Worldwide. For example, data may be disaggregated by business segment, or region. Complex data, reported in standardized XBRL format, can be easily understood and used, because of the underlying structure in which it is reported. An AI platform that does not immediately understand all the metadata associated with every fact used to fuel its outcomes, will not produce meaningful and accurate results.

**How data standards work**

When financial or other data is prepared in XBRL format, consistent metadata about each fact is embedded within it. The standardization process starts with a digital dictionary of terms called a taxonomy, that contains the hierarchy, relationships, definitions, and labels to unequivocally define a fact. A fact like 3,405 from the balance sheet example above would be defined by the term "Cash and Cash Equivalents, at Carrying Value" as shown on the figure below in the US GAAP Financial Reporting Taxonomy. The taxonomy defines the definition, label, authoritative references, and various properties (balance type, period type, data type, etc.) of the concept "Cash and Cash Equivalents, at Carrying Value", and also defines the relationship of the concept...
to other concepts that could be reported on a balance sheet. Relationships define how the data is presented (as shown on the left side of the figure below), and may also explain how the concepts are added or subtracted, for example, the fact that Cash and Cash Equivalents is added to other items to make up Assets.

When the balance sheet data is reported in an XBRL report (instance document), the reporting entity includes other needed metadata such as the units used (US dollars or some other currency), the time period, scale (millions, thousands), and the name of the reporting entity. Again, this metadata is embedded in the fact so that when it is used in an AI platform, it is completely unambiguous.

The figure below shows the same 3,504 fact, as it appears in the Hilton Worldwide financial statement posted on the SEC website. The balance sheet is shown in an open source viewer the SEC has created which allows viewers to see the embedded metadata in each reported fact by clicking on any of the facts highlighted in red (incidentally, the SEC’s open source viewer can be leveraged by any other regulator or commercial entity that wishes to show the XBRL embedded tags in an XBRL document). The reporting entity specifies that the taxonomy concept “Cash and Cash Equivalents at Carrying Value” is the appropriate concept and then adds information to explain that the fact represents the period as of 9/30/2020, is in US dollars, has a debit balance,
and a positive sign. This metadata appears in the gray popup box in the SEC viewer because it is embedded in the data point.

Because the balance sheet is prepared in XBRL format, every fact highlighted in red can be extracted automatically and with complete understanding of the meaning of each value.

Artificial intelligence platforms are only as good as the data supporting them. An AI platform driven by consistent, structured, unambiguous data formatted in the open, nonproprietary XBRL data standard will produce more accurate, useful outcomes.
We appreciate the opportunity to provide input to this important topic. I am available to respond to questions or to discuss further how the XBRL standard is used today and how it can be leveraged to ensure good quality data in AI platforms. Please feel free to contact us if you have any follow up questions or would like to discuss. I can be reached by emailing campbell.pryde@xbrl.us or at (917) 582-6159.

Respectfully,

Campbell Pryde,
President and CEO