



1211 Avenue of the Americas  
19<sup>th</sup> Floor  
New York, NY 10036  
Phone: (202) 448-1985  
Fax: (866) 516-6923

February 22, 2021

Regina Kearney  
The Office of Management and Budget  
Attn: OFFM  
725 17th Street NW  
Washington, DC 20503

Dear Ms. Kearney:

RE: Office of Management and Budget - Proposed Designation of Databases for Treasury's Working System Under the Do Not Pay Initiative

We support the efforts of the Office of Management and Budget (OMB) to aid federal agencies in their review of payments and awards to guard against improper payments through the DoNotPay (DNP) Initiative. We encourage OMB to adopt data standards for the information provided through DNP to improve the efficiency and timeliness of data access through this important program.

The U.S. government collects data from millions of entities including individuals, businesses, and governments. Leveraging this vast pool of information to protect against underpayments and overpayments, mistakes, and potentially fraud, is an important and worthwhile initiative, not only for the U.S. government but for the various entities entitled to payment. The Treasury proposal to include an additional 12 databases in the DNP web portal, could increase its value by providing more resources through which agencies can evaluate, track, and monitor the individuals and organizations receiving payment, and the payments being made.

XBRL US is a nonprofit standards organization. Our members include public companies, accounting firms, software, data, and service providers, as well as other nonprofits and standards organizations. The mission of XBRL US is 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 XBRL specification, which is a free and open data standard widely used around the world for reporting by public and private companies, as well as government entities.

As a data standards consortium, we recognize that the DNP web portal can only be effective if the data it provides to agencies is clean, accurate, consistent, and current. We also understand the challenges of collecting data, and in particular, of combining multiple sources of information.

The Treasury plan to add 12 more databases into DNP highlights the need, and poses an opportunity, to vastly increase the efficiency of the system through data standards.

### **Making data sources interoperable**

Much of the data reported through the various databases comprising the current and proposed DNP initiative is collected and made available in different formats. For example, the Census Bureau's Federal Audit Clearinghouse collects data from grantees in PDF files, and then requires the grantees to key some of the data already provided in the PDF into an online form. Once the data is received, it is made available to data users in the original PDF, and in partial datasets through text files and spreadsheets.

Some agencies collect data in a variety of forms such as PDF, spreadsheets, or online fill-in forms, then painstakingly extracts the data into their own custom relational database, and allow users to pull data through APIs or other querying tools they build. While this improves the usability of the data, the highly manual processing means that the data may be prone to error, and very likely outdated by the time it is available for use. The U.S. Census makes various APIs<sup>1</sup> available to access its data, but the most current data available is two years old. Data that is two years old by the time it is available, is significantly less valuable, particularly for the purposes of the DNP initiative. Plus, if the reporting requirements for a dataset changes, the government agency is tasked with costly updates to APIs and/or querying and extraction tools to fit the data changes.

Data submitted to the Securities and Exchange Commission (SEC) however, which is one of the proposed 12 new data sources, is prepared and submitted to the SEC in XBRL format. XBRL is an open, nonproprietary data standard that is widely used around the world and is required for reporting to certain U.S. regulators: banks reporting to the Federal Deposit Insurance Corporation (FDIC); public companies, mutual funds, credit rating agencies, reporting to the SEC; and public utilities, reporting to the Federal Energy Regulatory Commission (FERC). These reporting entities prepare and submit different kinds of data to these regulators and yet, their data is interoperable because it is prepared in the same structured format. "Interoperable" means that a data user can leverage the same tools to create, process, extract, and analyze any of these data sources. Datasets from the SEC, FERC, and FDIC can be combined and searched on simultaneously with minimal manipulation of the raw data to consolidate it or make it comparable. Interoperability should be an important aspect of data available through the DNP, as it needs to bring together multiple datasets from disparate agencies.

Adopting a standardized approach like that used today by the SEC, the FDIC, and the FERC, would vastly increase the efficiencies and usefulness of data reported through the DNP. If the various DNP databases contained data structured in the same fashion, the data reported would be more efficient, easy to access, and substantially less costly to maintain.

---

<sup>1</sup> United States Census Available APIs: <https://www.census.gov/data/developers/data-sets.html>

## Can data from different agencies be standardized?

The structured data standard (XBRL) is designed to handle the complexities of financial and other data types, and to manage complex dimensionality in data, if needed. It establishes structure around a reported fact to clearly define it, such that it can be read by a machine with no manual vetting or processing. Processing data from XBRL files is faster, more accurate, and cheaper than processing data from other file types.

Morningstar's Global Director of Equity stated in a recent video<sup>2</sup>, *"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..."*

Speed and accuracy are critical features of information used to assess whether a federal payment is going to the appropriate individual or organization. Many of the data sources comprising the current and proposed DNP have data that is financial in nature, for example data collected by the Internal Revenue Service (IRS) and the Census Bureau. Other proposed data sources contain different kinds of numeric and identifier data such as the United States Postal Service and the Veterans Affairs (VA) Beneficiary Identification Records Locator Service. Structured data standards are designed to easily handle data of these types (and others such as text or even text blocks) and to render the data reported unambiguous and fully machine-readable.

## Is structured data the right approach?

Introducing data standards into the various DNP sources would reduce the cost of data preparation, processing, and analysis, and would provide agencies using it more consistent, timely data in machine-readable format. Given that the SEC already produces a significant amount of data in XBRL format today, other DNP sources could be adapted to leverage the same XBRL standard. Adapting the data sources that are primarily financial in nature (like IRS and Census) first, would be an effective means to start. Other sources that provide text or other kinds of numeric data could follow.

This approach would not only be beneficial to the DNP program but would aid in reducing cost and improving efficiencies for the reporting agencies as well.

## Does it really work?

An approach to government data collection has been taken in non-US markets that has resulted in significant savings across business and government. In 2010, Australia launched a program to standardize business to government reporting (called Standard Business Reporting - SBR<sup>3</sup>). While implementation required a clear vision, patience, and upfront work, as of 2016, that program was estimated to save Australian businesses and government \$1.1 billion annually<sup>4</sup>. Similar

---

<sup>2</sup> Video "XBRL for Analysts and Investors" <https://xbrl.us/news/analyst-video/>

<sup>3</sup> Australia Standard Business Reporting: <https://www.sbr.gov.au/>

<sup>4</sup> Itnews "Tax office claims \$1 billion in savings from SBR": <https://www.itnews.com.au/news/tax-office-claims-1-billion-in-savings-from-sbr-432460>

programs are in place or underway in other countries including the Netherlands. Translated to the American economy, which has a GDP 10 times the size of Australia, the Standard Business Reporting approach has the potential to create \$11 billion in savings across the US economy.

The XBRL standard is a mature, widely used standard. As noted by XBRL International<sup>5</sup>, more than 180 digital reporting mandates exist across more than 70 countries, requiring digital reporting using XBRL by more than 10 million entities, including private and public companies, as well as governments. This year, the European Securities Markets Authority (ESMA) begins its mandate for XBRL reporting by public companies in every European Union country, plus the United Kingdom. More and more data is reported in standardized, structured format, both within the U.S. and in non-US markets. The more a standard is used, the easier it becomes to adopt because of the ubiquity of tools available to prepare, process, extract, and query using the standard.

The U.S. government already has a strong starting point with the SEC implementation of XBRL in the EDGAR System. We recommend building on that with the other DNP sources to give agencies the best tools possible to manage critical payments. Not only would this approach improve the efficiencies of the DoNotPay Initiative, it would result in substantial savings and improved data access across all government agencies adopting it.

We appreciate the opportunity to comment on this important topic. We urge the Commission to consider how data standards would improve the effectiveness of the DNP Initiative for government agencies and for organizations and individuals who received government payments. Adopting a single data standard will result in much greater benefits and we would be happy to discuss this further.

Please contact me with any questions. I can be reached at [campbell.pryde@xbrl.us](mailto:campbell.pryde@xbrl.us) or by phone at (917) 582-6159.

Regards,

A handwritten signature in black ink, appearing to read 'C. Pryde', is positioned above the typed name.

Campbell Pryde,  
President and CEO

---

<sup>5</sup> XBRL International project directory: <https://www.xbrl.org/the-standard/why/xbrl-project-directory/>