Re: BCBS D604 - Consultative Document Machine-readable Pillar 3 disclosure

We welcome the efforts of the BIS and BCBS to standardise and harmonise the technical aspects of Pillar 3 reporting. We hope that this work paves the way for future efforts in this area and we believe that the proposals will bring immense value to both the financial system stakeholders (banks, central banks, supervisors) as well as the wider public, given the nature of these disclosures. Please see responses to the consultation questions below:

Q1. What are your views on the scope of the requirement for machine-readable Pillar 3 disclosure, in particular the proposed initial focus on quantitative disclosure regarding the templates marked as "Maybe" in Table 4?

- Either way, this will become a project for the banks to implement. If there is a high probability of the "Maybe" templates becoming a "Yes" in the near future, then best to include them now rather than creating another potential project in the future.

Q2. Do you have any comments on the technical data formats and standards, the API structure and the general concept of the taxonomy proposed for quantitative machine-readable Pillar 3 disclosure? Are there other formats and solutions that should be considered?

- The best data format of the proposed 4 (JSON, CSV, XBRL, SDMX) is JSON by a significant margin.
- CSV is a lightweight, flat format but is not really standardised and struggles with nested data (lists) and data that may contain the character that is used to create the "separated values" whether that is a comma, a tab a pipe etc.
- Despite being mandated by many financial regulators for decades, XBRL and SDMX have no popular open source libraries or cost effective software applications.
- Additionally, XBRL/SDMX are not actually data formats. They are document formats (like HTML). The difference between a document format and a data format is that a document provides the data and describes it, whereas a data format just provides the data. Sending data in a document format is the equivalent of a database query returning data in HTML rather than just the figures. Using a document format for transmitting data between two trusted parties creates unnecessary bloat to the file size – to the extent that some XBRL files can be several gigabytes in size and need to go through zip compression to actually be sent now. This difference is self-evident in the examples given in Annex 3 of the consultation document.
- Using a document format for data actually makes the data less-human readable and harder to process upon receipt by the receiving party. Consider that all banks and central banks have to pay for an XBRL software rather than being able to use one of the thousands of free or open source JSON tools. Even Excel can open a complex JSON now (Open Microsoft Excel, go to the Data ribbon tab, and select Get Data -> From File -> From JSON.)
- It is unnecessary to tag and describe each data element in the level of detail XBRL requires when transmitting data. Both parties sending and receiving Pillar 3 data will understand agreed reference ids like "km1.row_030.col_020". There may be some additional metadata at the template level required, such as the currency, but this is minimal compared to what something like XBRL requires. If further information is required about the specific details of that datapoint, what it is called, how it is described, that information can still be made accessible elsewhere, but it is not data the bank should send, because the central bank *already knows* that information (what "km1.row_030.col_020" refers to), it is redundant.

- Using something like XBRL-CSV or SDMX-CSV to attempt to transform those document formats into a data format relevant for the modern age, is kind of like the worst of both worlds. It has all the drawbacks of those document formats with all the drawbacks of CSV.
- Finally, it is worth stating the obvious: JSON is the de-facto data format of all modern IT systems, choosing anything else will be handicapping this project before it begins. Any and every developer in the world is familiar with JSON, from new graduates software programmers to retiring systems administrators. In contrast, there are only a handful of people in the world familiar with and currently developing XBRL/SDMX software.

Q3. Are there formats other than PDF that should be considered for human-readable disclosure?

- PDF is a great *output format*, but it is no longer the optimal format for presentation of information for analysis due the evolution of how we present and consume information over the past 25 years. PDF is essentially an image that can be printed to paper consistently and reliably to ensure a harmonised user experience to all readers. The fact that this project is being proposed is evidence that current Pillar 3 reports in PDF format are not satisfactory.
- There are no longer output formats. PDF was often the end of a chain of data preparation and presentation. But modern systems, analytics and now AI means that there are only data consumption circles, not chains. So all human-readable formats must also be machine-readable.
- Today, nobody prints things to paper (it is in fact discouraged environmentally). The primary form of consumption of documents and data is HTML, whether that is in a browser, on the phone, in an HTML-enabled email client etc. Even a JSON response of data is never viewed as JSON but is almost always displayed in an HTML table. The tooling, support, compatibility, functionality for HTML has evolved exponentially over the past 25 years as well (due to demand). HTMLs can also be readily converted to PDF via many open source tools or even simply printing to PDF from the browser.
- An objection might be that HTML in its raw form (with tags) is not so human-readable. We must remember that PDF is also not in its raw form and its raw form is much less human readable. It should also be noted that PDF is the “presentation layer” of this raw form, just like a browser is for HTML. PDF viewers are also often proprietary, and not cross-compatible, intentionally (like Adobe).
- In addition, HMTL, due to its popularity has spawned numerous, open source and widely used Markup Lanuages, which can be written and read easily in any simple text editor, and auto-convert into HTML, often natively by many applications. The most prominent example of this is Markdown, invented in 2004 and formally standardised (https://en.wikipedia.org/wiki/Markdown#Standardization) by the Internet Engineering Task Force in RFC 7763 (https://www.rfc-editor.org/rfc/rfc7763).

Q4. In your view, which are the main operational benefits and challenges that this project would bring to banks? Would you see any other positive or negative impacts on your current disclosure process?

- The primary challenge is data conversion and implementation and this will be the primary push back. And the primary driver for that challenge is choice of format. A simple JSON format of template, row, column identifiers and a number in a pre agreed schema format will be infinitely easier to implement by any junior developer (a free, open source solution may also spring up). Whereas choosing something like XBRL or SDMX will immediately trigger procurement, costs, specialists, new systems integrations etc.

- Once implemented, the benefits are obvious as it makes data more accessible, transparent and comparable.

Q5. Do you believe the proposed effective date would provide sufficient time for implementation of machine-readable Pillar 3 disclosure? Would smaller internationally active banks need additional time?

- Not in JSON, yes in XBRL/SDMX for the reasons stated above.
- I am also confident that if the format is described in simple JSON, an open source solution or library will be developed by one of the community members given the limited competitive scope and widespread applicability of this requirement.

Q6. How useful would data users consider a global database on the BCBS's website? Would visualisation tools and industry aggregates make a global repository meaningfully more useful?

- Yes, absolutely. Having public disclosure information (Pillar 3 reports) become more machine-accessible is enormously beneficial. In particular, with modern AI tools, this information and analysis of it can be more widespread and can only enhance the original Pillar 3 guiding principles (https://www.bis.org/bcbs/publ/d309.pdf p.3). In particular, the principles of the information being readily accessible and comparable across banks.

Kind regards,
Murat Abur

Chief Information and Architecture Officer
Suade Labs