



24 June 2022

Alison Bliss  
General Manager  
Australian Banking Association  
1 Martin Place, Sydney, NSW, 2000

Dear Ms Bliss,

## Direction for data collections

The Australian Banking Association (**ABA**) and its members welcome the Australian Prudential Regulation Authority's (**APRA**) Directions for Data Collections and its associated five-year road map. The outline of APRA's vision and the reasoning underpinning its strategy has allowed banks to start planning for and resourcing what is to be a major regulatory change, requiring significant investment.

ABA members are now looking for the next level of detail on the future data collection, so they can start their own projects to ensure delivering to APRA's timetables in the most efficient manner possible. This submission outlines those areas in which ABA members are seeking that additional detail. These fall into:

- **Planning:** Staging new collections and phasing out older forms
- **Costs and benefits to industry:** Implementation, governance, and cross-agency data
- **Operating model:** Transparency over APRA's data collection operating model
- **Communication and engagement:** Ways of working with industry

Omitted from this document is commentary on any costings to industry or alignment of the approach to organisational data strategy. These will be considered in the individual submissions made by ABA members.

## About the ABA

The Australian Banking Association advocates for a strong, competitive and innovative banking industry that delivers excellent and equitable outcomes for customers. We promote and encourage policies that improve banking services for all Australians, through advocacy, research, policy expertise and thought leadership.



## Planning and project management

ABA members welcome APRA's overview of their intended data collection strategy as outlined in the Directions for Data Collections discussion paper (**DDC discussion paper**). Banks are now looking for a further level of detail to guide their project planning, investment, and resourcing of the implementation over the coming five years.

### From PET forms to PET collections

The DDC discussion paper describes APRA's motivation to transition all reporting forms to APRA Connect as driven by decommissioning the current data collection tool, D2A. It also outlines the intention to develop interim reporting collections, a comprehensive credit collection, a comprehensive liabilities collection and "several simpler, smaller collections that focus on areas such as market risk, interest rate risk in the banking book and financial statements" (hereafter collectively referred to as '**the collections**').

From this description ABA members understand that APRA intends to merge all economic & financial statistics (EFS) and prudential reports into two large data collections (i.e., the comprehensive credit collection and the comprehensive liabilities collection) with some smaller collections accounting for the remaining collections which do not fit into the two larger ones.

What is not well understood is whether (and how) all prudential concepts will be linked to EFS concepts within each collection, or whether there will be multiple aspects to each collection which will avoid the need for linking disparate concepts.

Another aspect of the changes not yet fully understood is whether APRA intends to retire all Plain English Taxonomy reporting forms (**PET forms**) and transition them into data collections, or whether there will remain some PET forms that are collected through APRA Connect. If all forms are to be granular, questions remain as to how this will be done for concepts such as non-financial risk and profit and loss.

### Data mapping

To make APRA's intentions clear and to help banks with their own project management, we encourage APRA to map out the high-level structure of the collections as early as possible. This exercise would begin with mapping the data fields in the PET forms to the data fields in each collection. The exercise should also involve assessing the data requirements from peer agencies and integrating them into the data mapping. For example, the Survey of International Investment, collected by the Australian Bureau of Statistics, could be integrated into the APRA data collection during this process.

The data models associated with each collection would *not* be developed in this exercise. They would be developed sequentially over the coming five years. The mapping exercise would be the identification of which PET forms would be retired and at what time. We consider this the most efficient way to transition the current form-based reporting to the collections-based reporting as it allows banks to map out their own resourcing needs over the coming years. Clear scope and definition of each collection and the existing PET forms they will be replacing.

Another outcome would be a framework which represents all collection across all phases, including those requirements from peer agencies. It would include clearly defined data structures for each collection with specifications about how the structures relate to each other. It would include clear timelines, dependencies and a framework for each collection.

### Sequencing of collections

As described above, the mapping exercise should produce a document which outlines the sequencing of the collections. It is likely that each reporting entity will have their own preference as to how to phase the collections and so ABA members, as a collective, have no specific comment on the matter in this submission. We do encourage APRA to work with all reporting entities after this consultation to develop



the sequencing in a way that works with the needs of most regulated entities, and to consider delaying collections impacted by the APS 330 consultation to ensure adequate time is given to develop this.

## Artefacts

As a final stage in the development of the collections, all ADIs should be given the opportunity to consult on artefacts associated with each collection, including:

- **Data dictionary** including definition, data type, mandatory values
- **Validation rules** (e.g. “Cannot be well secured with zero collateral”)
- **Taxonomies** which identify how each granular data attribute will be used to aggregate or group at a form / output level. For example, where replacing an existing aggregated form with granular data requires a mapping between the two grains of data (e.g.: book value = drawn + undrawn amount)

## Costs and governance

Costs to banks are difficult to measure given the current high-level guidance of the reporting changes. Banks will be required to change their data architecture to source, test, review and supply large volumes of data. They will require increased governance over the datasets and have little understanding of how to do so at this point in time.

In the current form-based reporting banks can apply governance to every data concept that was provided to APRA. However, given there are many ways in which the data will be able to be examined through the collections, it will not be possible to apply governance to every data concept. Banks will need to map their own data to APRA’s data dictionary and build an aggregated reporting layer on top of the granular data to test whether it satisfies APRA’s reporting expectations and potential queries. However, all of this is dependent on understanding how APRA will use the data.

We encourage APRA to separately consult on auditing and quality control processes. We consider this work to be as complicated and time consuming as the development of the collections themselves. Delaying this consultation may prevent APRA from meeting the deadlines as set out in the roadmap.

## Operating models

It would be useful for banks to understand the various operating models which guide the way APRA conducts its work, including APRA’s own internal operating model as well as the way APRA interacts with agencies.

### APRA internal operating model

ABA member banks would benefit from a clear understanding of how APRA teams work together to develop and analyse data reported by regulated entities. For example, to what extent and how does the APRA data team consult with APRA policy teams during development? The benefit of industry knowing this information is that they can develop operating models which mimic APRA’s. That is, if it is clear at which point policy specialists are consulted with in APRA, banks can ensure they have the equivalent staff review their contributions to the development of collections at the same point.

### APRA / Agencies operating model

Further, given the intention for APRA to become the key data collector in government (an intention that is supported by the industry) it is useful to understand how APRA intends to work with these agencies. Questions arise about the three stages of working with agencies: development of collections, provision of existing data to agencies, and management of queries.

Outlining a clear operating model between APRA, the agencies and banks from the outset will allow banks to understand what to expect from agencies. This would provide clear parameters for expected collections, including anticipating the volume of requests, what quality is expected, and how to apply governance to the data.



## Development of collections

During the development of any new collections, we encourage APRA to collaborate with other agencies in a way which maintains an open and meaningful conversation with banks. A high trust relationship is necessary during the any development phase which involves the sharing of bank data to regulators. This is due to the data not having the same level of governance as would exist in formal data sharing.

Therefore, where agencies are involved in the development of collections it is essential they understand potential data quality issues and are committed to using the data only for the purposes of assessing the development of a collection, and not assessing the entities providing the data. Without this certainty, there is the potential to damage the goodwill on which collaborative development of collections is founded.

It is also useful for banks to understand how APRA intends to manage the data needs and wants of agencies during the development of any collections. Experience of the development of EFS was that APRA played a facilitative role in the development of collections on behalf of the agencies but did not play an authoritative role. This resulted in some concepts which were difficult for banks to capture from their own systems.

With multiple consumer agencies using the granular data there is the risk single data points will not meet all the different Agency requirements. This has the potential to either result in duplicated data points to meet different definitions or even agencies using data with a misunderstanding of a data point definition. Again, it would be useful for APRA to have transparent guidelines over how these potential conflicts will be managed.

During the development of the collections and across the transition, there may be opportunity for APRA to reduce burden on industry further by reviewing the current ad-hoc reporting that is provided to supplement the aggregated APRA return forms. This reassessment may highlight gaps in the granular data requirements and give APRA opportunity to capture what is really needed in a more cohesive manner.

## Provision of existing data to agencies

Banks would benefit from certainty as to how data which is collected by APRA will be distributed to the agencies, and the extent to which banks will have oversight of what is shared. As discussed earlier, the many permutations by which the data can be examined will make it difficult for banks to apply governance to all aspects of the data. Therefore, distributing the data to other agencies without initial oversight from the reporting entities may cause unnecessary complications in regulatory supervision.

We suggest that APRA develop a model by which reporting entities are first made aware of what data is being shared with agencies before it is shared, so they can provide appropriate oversight as well as identify and correct any anomalies. It will also allow reporting entities to ensure they understand and can explain the context for any data before receiving queries from agencies. This may come in the form of a published data matrix with details of what data is being shared with what peer agencies and when.

## Query and resubmission management

As part of the forward planning of data collection we recommend APRA review how reporting entities are queried about data submissions. Responding to queries creates a large burden on banks as often:

- similar queries are asked by multiple agencies
- queries are often raised where there are only slight differences
- only a short amount of time is given to respond to queries.

Banks receive queries and resubmission requests from multiple agencies that oftentimes overlap, and / or queries with logical answers that could be answered before the query gets to the bank. If there was a single point of contact who could manage the queries and identify patterns that provide responses to agencies on a regular basis, this would minimise a huge workload for reporting teams. This would free up more resources to apply to the change management needed for implementation of the new APRA data collections. We propose that with APRA as the primary collector of data for government agencies, they should also be the primary querant of any data anomalies or outliers.



Consideration could also be given to how resubmissions are managed, including providing a greater threshold for which resubmissions are required. For example, banks resubmit data where there is only minor discrepancies in line items due to larger changes in the 'total' values into which these line items feed. However, the line item changes do not meet the threshold requirements for resubmission. We suggest this requirement be reviewed.

## Communication and engagement

APRA's methods of engaging with stakeholders when developing and implementing reporting changes has become more inclusive in recent years. Early engagement with reporting industries and multilateral engagement through industry associations has benefited ABA members. There are some changes we encourage that would benefit all stakeholders including clear communication and detailed taxonomies. As well as greater gaps between final reporting standards and first data submissions.

### Sharing data insights

Little is understood about how APRA will (and can) use the data it collects. Once APRA has access to the substantial amount of banking industry data many insights can be generated, but what these insights are and how they will be used is not known by reporting entities. APRA has stated that the COVID19 collections have benefited APRA supervision of the financial sector but have not shared what insights they have derived from the data.

### Clear and prompt communication

Clear communication about all reporting changes is useful for all parts of industry. While flexibility is needed for development of the data collections, banks are still required to plan and resource future changes and make their own bids for investment. Communicating any of APRA's expectations with internal bank stakeholders is difficult where there is no formal communication on the matter.

For example, the Comprehensive Credit Collection has evolved out of the development of, and has replaced, the Strategic ARS 220.0. However, there is nowhere where APRA has made a public comment on this matter. Consequently, some industry participants are still expecting to implement the Strategic ARS 220.0 as described by APRA in their letter to industry of June 8, 2021.

One way of keeping industry updated on the development of the collections is for APRA to minute industry meetings and distribute / publish the minutes. Although not all agenda items and actions would need to be published, this could be a mechanism by which all of industry could be kept abreast of any developments.

### Timelines for implementation and reporting queries

Experience with interim ARS 220.0 reporting was that the tight timeframe between the publication of the standard and the implementation meant there was very little time to test the reporting solutions, identify anomalies and query APRA on definitions. There were also long periods between which banks would raise queries and APRA would respond.

We suggest there are substantial periods between the development of reporting standards (or collection taxonomies) and the first submissions. ABA members find a period of 18 months to be appropriate as this gives them enough time to source data, build data infrastructure to capture, store and transmit data, and to test the solution which has been developed. For smaller reporting changes, we encourage APRA to provide a lead time of 12 months, as this again requires the sourcing of data, alterations to data infrastructure and testing the changes.

Further, banks need a mechanism by which reporting questions can be answered as quickly as possible. It would also be beneficial for APRA to commit to timeframes in which they will respond to queries.