



# Data Standards Body

## Non-functional Requirements (NFR) Consultative Group

### Minutes of the Meeting

*Date:* Thursday 27 March 2024

*Location:* Held remotely, via MS Teams

*Time:* 14:00 to 16:00

*Meeting:* Meeting # 2

## Attendees

### Participant Members

---

James Bligh, DSB Chair  
Andrew Ferris, AGL  
Dhananjay Gourshettiwar, Westpac  
John Adshead, AEMO

Jon Denly Basiq  
Joseph Lucas, ANZ  
Julian Luton, CBA  
Mark Wallis, Skript

### Observers

---

Elizabeth Arnold, DSB  
Hemang Rathod, DSB  
Nils Berg, DSB

Mark Verstege, DSB  
Terri McLachlan, DSB

### Apologies

---

Jim Basey, Basiq



## Chair Introduction

The Chair welcomed everyone to the second NFR Consultative Group meeting.

The Chair noted that since the last meeting he has resigned from the Data Standards Body (DSB) and he will be handing over his responsibilities to Mark Verstege, the DSB's Lead Architect. This will include Chairing the consultive group meetings.

### Minutes

The Chair thanked the group for their comments on the minutes from 22 February 2024 meeting. As per the Terms of Reference, the minutes will be made publicly available. The Chair sought any last-minute feedback prior to approving.

The Minutes were formally accepted.

### Action Items

The Chair noted that a number of Action Items arose from the last meeting. An update follows:

1. DSB to create a collaboration space for the group

Teams had been set up as a collaboration space for the group, and if anyone had any issues accessing, to let the DSB know.

2. DSB to invite Xero to a future meeting

The DSB thanked Skript for the Xero contact. Initial contact had been made with a follow up call scheduled for mid-April to discuss how they might want to engage with the group.

All other Action Items had been completed.

## Governance

### Terms of Reference

The Chair noted that the Terms of Reference (TOR) had been approved by the DSAC and the Chair of the DSB. The TOR had also been circulated to the group for further feedback and/or concerns.

As there was no further feedback, the TOR were formally accepted and will be made publicly available.

### Code of Conduct

The Chair noted that the main purpose of the Code of Conduct (CoC) was to cover the sharing of sensitive information which would include metric data from participants or potentially ACCC so that the group can do forecasting etc.

The Chair was seeking input from the group on the draft CoC and whether it was acceptable and met individual and organisational needs.

One member noted that they are comfortable with the CoC and called out explicitly being permitted to share the confidential discussions within their organisation on a need-to-know basis as they are a fairly small team.



One member noted that were also supportive and had no concerns.

One member noted that they are comfortable with the Chatham House Rules as far as attributability. They are also still working through what data they are comfortable sharing with the group.

The member also noted that they like the concept of data shared within the group and also the option for more sensitive data to be shared directly with the DSB and/or the ACCC. However, they will need sufficient time to get approvals at multiple levels in place within their organisation.

The DSB noted that if the internal processes are not getting easier during the 6-month trial, we can take that as a finding from the experiment and present it to DSAC to highlight and/or suggest some fast-track processes to service the CDR for consideration.

The Chair noted that as there was no further feedback, the CoC were formally accepted. He did note that if they are not fit for purpose moving forward, they can be revisited in the future.

The Chair handed over to Mark Verstege – the new Chair of the NFR Consultative Group to Chair the remainder of the meeting.

## Forecasting

The Chair noted that one of the reasons of the NFR Consultative Group was to forecast and prepare for what is existing, for organic growth and to run some “what if” scenarios. For example, what if we had a new entrant to the CDR who created a significant load and what would that look like (e.g., Xero or MYOB) or what if screen scraping was banned. We need to be evidence based and plan into the future.

Nils Berge from the DSB noted that he looked at ACCC data and based on that considered the forecasting moving forward. Some highlights following:

- Data used in the forecasting includes data from the ACCC, the availability and invocation data which is feeding the public performance dashboard and data directly from the performance dashboard
- Forecasting includes the average availability per day for all data holders (DH), all sectors, both sectors and total invocations per day going back for 3 years.
- Based on the first basic projection, the linear projection is approx. 3,000,000 invocations per day based on the current traffic for all DHs and all sectors.
  - Projected to approx. 4,000,000 in the next year
  - Assuming margin of error, but with other platforms coming in or increased traffic, projecting approx. 7,000,000 over the next year
- Based on a simple trend line they also project approx. 4,000,000 invocations per day by next year.
- Under the current requirements, availability has not been affected by the increase and appears to have levelled out.

One member asked if an ADR is rejecting calls due to a throttle being reached, would that be picked up on availability and when DHs holders are starting to bump up against throttles. The DSB confirmed that this would be in the invocations.



Another member asked for data around when a DHs was shredding loads with Error 429 as the innovation count would go up.

The DSB provided a series of charts on invocations type (all time), invocation type (last 30 days) and invocation types (last 30 days – Big4) based on all brands. Data was provided on:

- high priority invocations
- unauthenticated invocations
- large payload invocations
- secondary invocations
- secondary DH errors
- low priority invocations
- unattended invocations
- primary DH errors
- large secondary invocations.

The charts showed the following:

- Unattended invocation percentage increased from 44% (all time) to 58% (last 30 days) to 66% (last 30 days – Big4).
- High Priority invocations decreased from 25% (all time) to 21% (last 30 days) to 16% (last 30 days – Big4)
- Low Priority invocations increased from 18% (all time) to 11% (last 30 days) to 15% (last 30 days – Big4)

One member was curious as to why the ratios between high priority and low priority were so wildly different. For example, bank # 1 was 2.5M to 9.3M and bank # 2 was 1.4M to 4.2M – very different ratios.

The DSB noted that bank # 1 was slightly higher on the error count at 0.94% whilst bank # 3 was 0.11%.

One member asked why bank # 3 was eight times higher (4,238,742) in terms of low priority invocation than the others? Other banks were for example bank # 1 = 934,187, bank # 2 = 590,861 and bank # 4 = 384,889. As this was an outlier, the number would need to be checked to understand the data behind it.

One member noted that a safe linear projection makes us feel comfortable but may give a false sense of security in terms of what's going to bite us i.e., a step change versus a liner change. For example, screen scraping which they believe is being used by small to medium sized businesses in banking and they are concerned about what may happen with load in the SME banking space.

The DSB asked the group for feedback on what scenarios should be modelled to present to the community? For example, a maximal projection at the upper limit, and what happens if screen scraping gets banned or gets made non-functional by MFA being applied to digital offerings.

One member agreed with screen scraping as they have much larger volumes on screen scraping than under CDR currently.

One member noted that it is good looking at the data within the CDR, but it would be beneficial to look at screen scraping and what's outside the CDR. How do we go about getting that data.



One member noted that they can get some numbers of what would happen if screen scraping was banned tomorrow, and everyone switched to see what the volumes would look like. They also noted that there has been a lot of friction for people trying to share their business accounts under CDR and they have stuck with screen scraping because it's easier.

One member noted that at least two of the big banks have manual paper-based processes that need to be completed to set up nominated reps because the volume is not there. This will be fixed when the load increases. They would like this problem addressed within this group forecasting how this manual process needs to be changed to automate now for those banks.

The DSB noted that there are mechanisms for businesses to share data through batch files with a charging method associated. For example, if an accounting platform comes in, the load won't be shifting from screen scraping because they don't use scraping, they use batch load with the banks which is a net new load from batch files which is another scenario and effectively more business usage oriented.

One member understands the bank reluctance as it is profitable which makes perfect business sense. There is also no indication for the removal of bank data fees.

Another member noted that they have been in discussions with an accounting platform around a post feed world or one where CDR plays a significant bigger role than it does at the moment, and it takes the load off bank feeds.

The DSB noted that it would be good to ask an accounting platform how they feel about moving away from feeds. It was noted that it is not ideal moving to replace bank fees with transaction APIs as it is a very efficient process vs inefficient, however it removes the fees.

The DSB noted they have been discussing internally a voluntary asynchronous mechanism in the CDR which could be the CDR facilitating a batch feed type scenario.

The DSB sought feedback on any potential scenarios where for consumption for energy data increasing in the near or distant future.

One member noted that one scenario is around isolate, jumping on board and starting to use CDR. Sales also continue to arrive through third party aggregators via established channels and everything else is just organic growth.

One member was curious if those types of product selection sites are not using the energy data, who are the customers for it?

One member noted that there are two providers one who represents 80% of their consents that they have on record and the other one has very moderate numbers. If the first provider came out with a vastly superior experience than what was currently offered, then perhaps you could see an increase.

One member noted that you may see a step change in usage if/when Public financial management systems (PFMS) that aggregate data from entity and suggest to customers a better deal.

One member noted that the chart presented was based on total invocations and it would be useful to have an alternative view where you have a peak TPS per DH view to show the thresholds as they are focused on authenticated traffic on a DH basis. We could then see how the utilisation of the current and the previous thresholds looks.



The DSB would be interested in seeing a similar breakdown for the big three retailers in energy to get a perspective of shared responsibility versus retail held data sets and to see where the growth is over time.

One member noted that they are seeing a lot of non-genuine traffic (i.e., genuine invocations, genuine customers but not genuine users of CDR) and the ADR are coming on board and doing a lot of testing internally. Filtering out that would be an impossible task and it does skew your view of what is happening in the CDR space in energy.

The DSB suggested a couple of things for discussion at the next meeting. The review of the DH side to understand some of the outliers or the peaks that we've seen and what sort of representations of the data are going to be useful. They also asked for other scenarios the group would like to discuss.

One member suggested a scenario around traffic thresholds for individual software products as they may cause a problem if you have a particularly large software product.

One member suggested a simple ratio of invocations to unique consents which might provide a picture as to whether we're seeing a ratio that is growing or fluctuating or out of step with expectation. This is around the energy sector and in particular around unwelcome repeats calls from ADRs.

## NFR Issues for discussion

Hemang Rathod from the DSB talked about a number of issues identified around NFR issues.

The DSB noted that these issues were sourced from change requests that have been raised and with energy in mind but after the conversations today, they believe they are sector agnostic.

A summary of the issues follows:

1. Data sharing arrangements with large number of accounts
  - Energy: C&I/large consumers in energy with 1500+ account with a single retailer
  - Single consent for data sharing with consumers who have a large number of accounts
  - Banking: ADR with large consumer base with large number of accounts or scenarios of mass migration and bring large number of customers with multiple accounts
  - Solutions for consideration: limit max number of accounts; limit max result set per endpoint; introduce a max number of servicePointIDs etc.
2. Low velocity, large volume data sets
  - Large data sets have a low velocity of change frequency
  - DHs have concerns where ADRs make multiple requests for the same data over a very short period
  - DSB make changes to the standards, that give DHs the ability to reject calls after a certain number
  - Issue with DHs having to police access to AEMO designated data, AEMO is better equipped to do this (for energy sector data sets)
  - Solutions for consideration: Add DH protections in standards; moving controls to SDH; introduce alternative patterns for sharing data sets
3. Customer Presence awareness for Energy Secondary Data Holder (SDH) (AEMO)



- Current NFRs call out different requirements for customers present vs non-present API calls
  - Current SDH has no means of determining if a call made is customer present or non-present
  - Mismatched performance expectations
  - Solutions for consideration: continue with AEMOs current approach; introduce a change to make AEMO customer aware
4. Alternative patterns for sharing of large volume low frequency changing data
- Pub-sub model (fire and forget)
  - Last update check API
  - Webhooks
  - Polling (not recommended)
  - Share recently updated/ new data only

One member wondering if there was an easy win to extend the concept of returning an Error 429 for something that falls outside of an NFR which is what we are already doing today.

Another member noted that from an ADR point of view, they can see an issue with using the same code without any extra information because they won't know the reason. They suggest using a "not before header".

Another member noted that the "not before headers" are mandated in the standards and are optional.

The DSB asked the group if there are particular models they use, for e.g., customer present or unattended or a combination of both for different uses cases for data?

One member noted that it depends on what the customer is using it for as they have lots of people using it for different use cases.

One member gives their customers the option of the frequency of when their data is refreshed and as an aggregator, they don't always have visibility over the full product end to end as they are acting as an intermediary with the data.

The DSB noted that a historical batch of 2 years of data could potentially suit a batch or async type of process where you will have the real time customer present and want to call the API and get their data from the last hour for example.

One member noted that they would like to discuss in more detail the "share recently updated/new data only" option. As an ADR this would have a significant effect on the amount of data they have to pull.

The DSB sought the feedback from the group around prioritising the issues raised, whether anything is missing and is there a particular order they would like to tackle them?

One member noted that the "large number of accounts" is a big issue, the "low velocity" is cached and therefore not a performance issue and the "customer presence awareness" is there lowest priority.



One member noted that they have a customer who has 400 sites eligible to be shared under the one sharing arrangements. They have accessed but not consented and they image the ADR will trigger a request for information for 400 sites all at once.

The DSB summarised the items for discussion at the next meeting:

- Asynchronous or event driven approaches
- Would we consider a voluntary pattern or a required pattern into the CDR?
- Single transaction event notification and to what extend they are practical

One member would be keen to look at changes from a business benefits lens perspective. For example, ways to mitigate a forecast increase in load by implementing a different pattern for returning data. What is the cost to implement and the benefit we expect to get from spending that time and money and does it make sense to go ahead with that.

The DSB noted that it is not their job to try to guess or calculate either the cost or the benefit of change but to take the feedback on topics into account with the change being incorporated into the standards and advise the Chair. This group is not determinative, but a good forum for the conversation to pre-stage conversations, get issues on the table but ultimately everything needs to go into a Decision Proposal for full public consultation.

The DSB also noted that it would be good for this group going forward to turn some of that ambiguity in a 'good citizen' and be more specific about expanding out that ADR requirement.

## Meeting Schedule

The Chair noted that the next meeting is scheduled for 24 April 2024 along with a forward view of the remaining meetings until the 17 July 2024.

## Any Other Business

The DSB sought feedback from the group about providing an update of the meeting to the Implementation Calls following a meeting for transparency. The DSB suggested posting a paragraph in the Teams channel summarising the meeting for the groups review, and which can be given as part of the stream update at the Implementation Call.

**ACTION:** DSB to post a summary of the meeting in Teams for the groups review for the purpose of providing an update at the weekly Implementation Call.

One member has concerns about perceptions outside of this forum of validity given there is only one energy retailer and one secondary DH attending. They would like to see more energy players in this group. It was noted that the current group put forward nominations for this group and were keen to attend.

The DSB noted that they can raise this at the next Australian Energy Council (AEC) meeting as it would be good to have larger DHs participating in this group. If a potential member was put forward, it was agreed to seek the Chair of the DSB advice in terms of whether to bring them into the group immediately or wait until the trial period ended (i.e., 6 months).

**ACTION:** DSB to speak to the AEC about a potential participant (larger DH) for the group.





**ACTION:** If potential participant identified, DSB to seek the DSBs Chair's advice on when they should join the group.

## Closing and Next Steps

The Chair noted that they will come back with further analysis of the data they presented at the meeting today. They also welcome feedback from the group around any thoughts on ways to present that data. For example, diving into some of the comparisons of invocations versus errors or rejections as well as understanding what is authenticated traffic versus unauthenticated.

The Chair noted that they will use Teams as the mechanism to take forward some of the priority discussions including asynchronous and event driven architecture and understanding that from a consumer and use case perspective.

One member volunteered to dig into their side and investigate ADRs asking for years' worth of transaction data and to see if it is an issue and report back to the group.

One member suggested that we have an agreed problem definition statement for what we're solving for.

The DSB noted that anything to do with sharing of large volumes of data with large number of accounts is a natural problem statement definition and the discussions around asynchronous or current real time versus batch historical and those solutions blend into solving that problem.

**ACTION:** DSB to draft up a problem definition statement for discussion at the next meeting

The Chair thanked the participants for attending the meeting.

The meeting closed at 15:53.