



Consumer
Data Right

Consumer Data Right

Data Quality Introduction & Scoping Workshop

Introduction of core concepts, problem space, challenges and long-term objectives.

August 2020

cdr.gov.au

Agenda



Defining the problem

1. What is data quality and what are the issues we will commonly face with the CDR?
2. Why does data quality matter?
3. What are the outcomes we want to achieve?

Introduction to Data Standards Conventions

1. What are conventions?
2. The role of conventions.
3. The relationship of conventions to standards.

Managing data issues

1. Overview of the process.
2. How will issues be identified and resolved.

Compliance & enforcement

1. What are the implications of data quality issues for compliance?
2. What role does enforcement play in data quality?

Next steps and other business

1. Summary of the workshop, next steps from here and any other business.

Why are we here?

—

Context:

- Initial four Data Holders launched product reference data (PRD) for banking on 1st February 2020
- Real-world consumer data sharing launched on 1st July 2020
- Non-major ADIs must launch PRD by 1st October 2020
- The CDR ecosystem is still emerging; consistency is still being tested and impacts understood
- Expect issues with interoperability, consistency and compliance
- Primary focus on PRD data; seeing growing use and review of PRD data
- Focus will shift to consumer data as ADR & DH base grows
- Success requires input and commitment from the community
- Ultimately, measured by consumer value



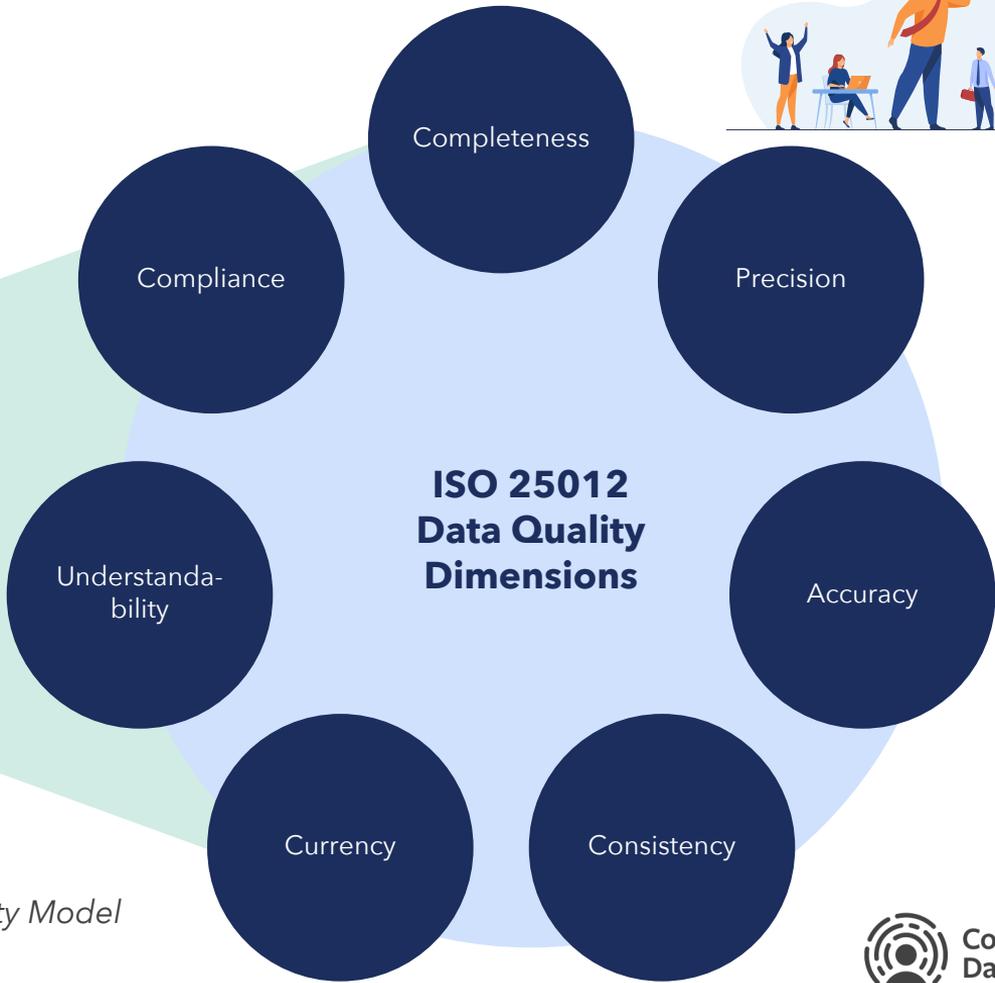
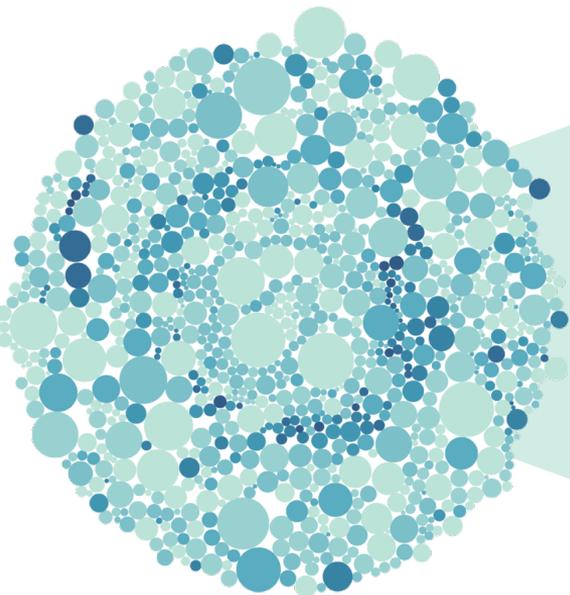
Objectives:

- Characterise the problem
- Describe the process to addressing data quality
- Gather feedback and input on approaches to resolving data quality issues as they arise

Out of scope:

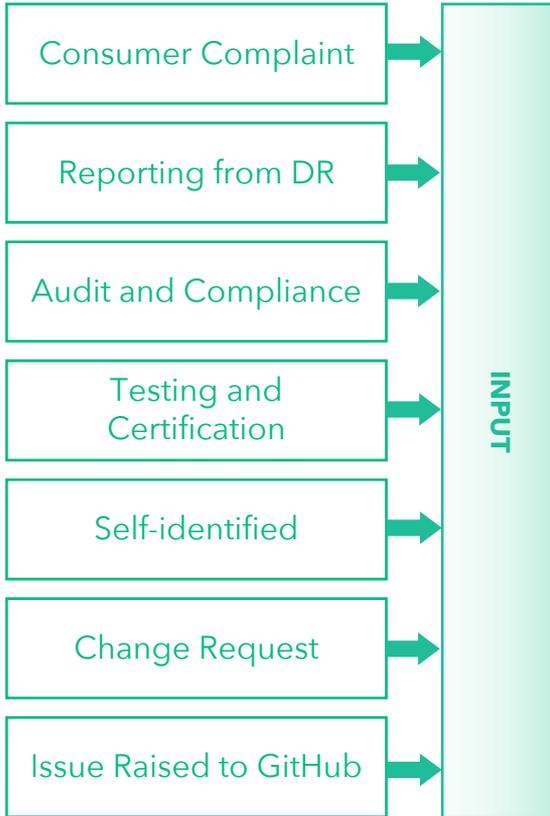
- Addressing syntactical correctness of the CDR payloads. This is considered a minimum ticket to play for DHs

What is data quality?



Source: Subset of the ISO/IEC 25012 Data Quality Model

Common issues with data quality in the CDR



Incorrect Source Data	<i>DH may present data correctly but the data itself is incorrect</i>
Deliberate/negligent non-compliance	<i>DH fails to provide required data</i>
Lack of Product Relationship	<i>Lack of linkage / association between products DHs offer in market versus the products customers own</i>

Individual Data Holders

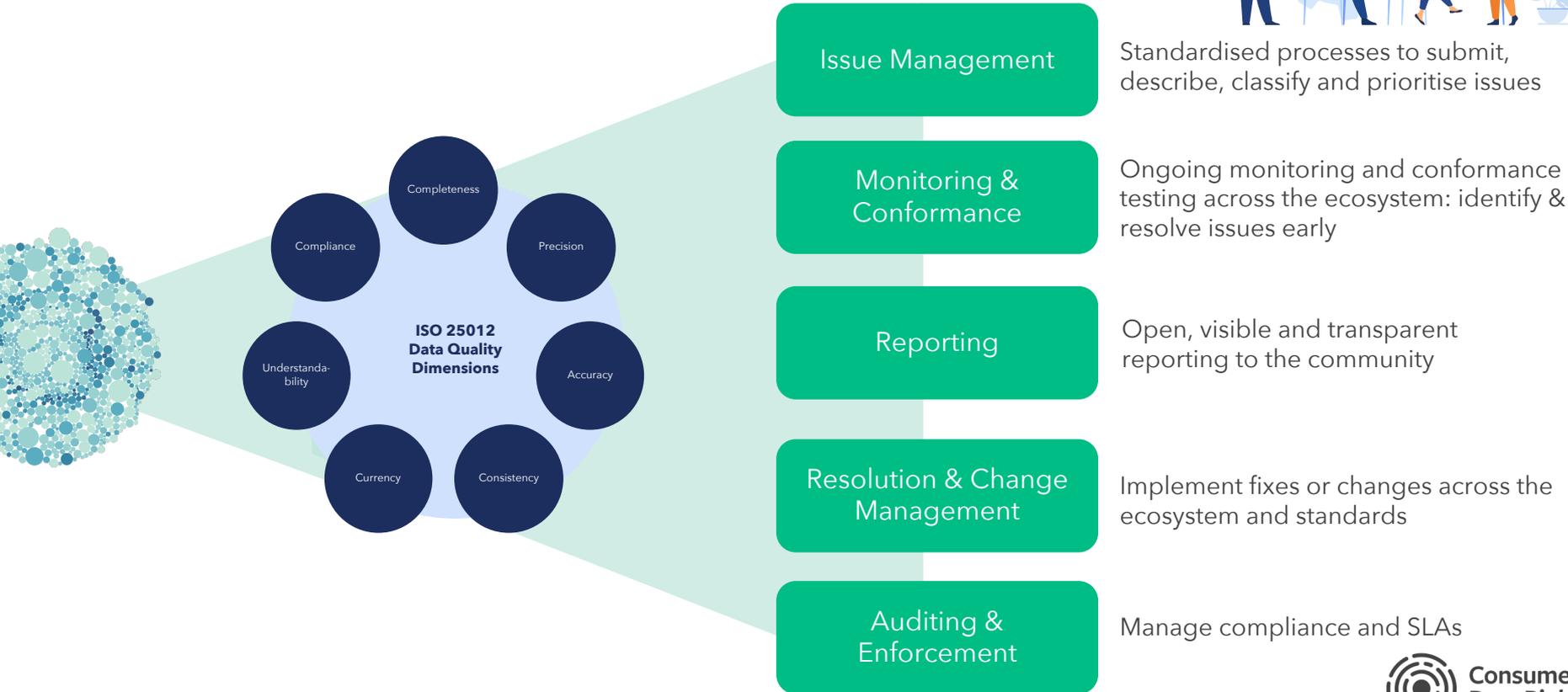
Implementation Error	<i>Software bug or failure to implement solution correctly</i>
Interpretation Error	<i>Failure to correctly interpret the standards correctly</i>

Many Data Holders

Ambiguous Standards	<i>DHs populate the payloads in differing ways that are all at least arguably compliant or consistent with the standards</i>
Overly Flexible Standards	<i>The standards provide too much flexibility for DHs, resulting in data that is too variable and/or to poorly structured to be useful to ADRs</i>
Incomplete Standards	<i>The standards are insufficiently expressive to capture details that DRs want/ need</i>

Standards

Key concepts for managing data quality



Motivate and support good behaviour



- Compliance is not just about enforcements. Need to motivate and move behaviour to the bottom of the pyramid



Introducing conventions



Conventions are a set of norms to describe generally agreed meanings and usage where there may be many objectively possible options and/or where there is insufficient knowledge to support fixed and definitive standards

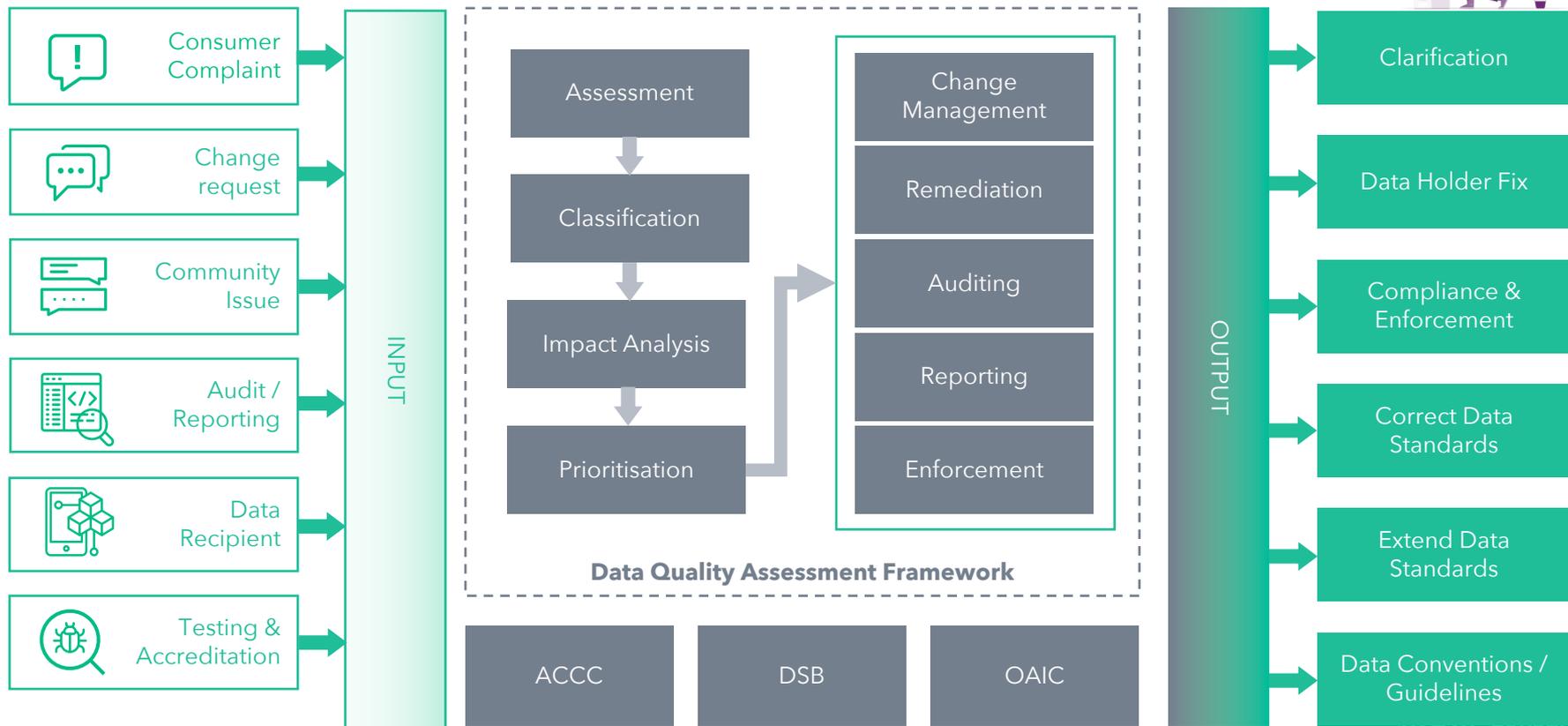
- Provide guidance on how to interpret the standards and how to apply existing data models to the CDR
- Describes the expectations on how things ought to look
- Fill the gap where semantics specific to a payload or data field aren't possible
- More flexible to change and adapt over time vs standards and rules
- Can adapt to market innovations and best practices more easily

Role of conventions



- Analogous to CX Guidelines vs CX Standards
- Complement the standards but aren't part of the standards
- Establish a reference external to the standards that documents "accepted conventions" for each contentious field
- There is only a *preference* that conventions be followed but in some circumstances they may become de facto obligations
- Less likely to be testable
- Allow standards to remain more stable; conventions that become widely accepted as standard practice may over time be ported into the standards themselves.

Managing the process for better data quality



Raising Data Quality issues



If you...

Then do this...

Link

Have identified a Data Quality issue that appears to be the result of non-compliance with the Rules or Standards by one or more identifiable Data Holders or Data Recipients.

Notify the ACCC. Preferably with documentation supporting your concern.

<https://www.cdr.gov.au/>

Think you may have identified a systemic Data Quality issue (i.e. an issue that appears across multiple parties) but realise that this could be a result of your misinterpretation or incomplete understanding of the Rules or Standards.

Submit a question via Zendesk OR post a clarification request via GitHub.

[ZenDesk Support Portal](#)

[GitHub Standards Maintenance](#)

Believe that you have identified a systemic Data Quality issue that you believe is best addressed by changing the Standards.

Submit a change request via GitHub.

[GitHub Standards Maintenance](#)

Believe that you have identified a systemic Data Quality issue but aren't sure how it should best be addressed.

Submit a clarification request via GitHub (clarifications may be made in the Standards or in the Conventions).

[GitHub Standards Maintenance](#)

CDR Rules: product data sharing

Overview

DSB Data Quality Workshop - 4 August 2020



Product data requests

- Data holders (DH) must provide a product data request service that enables any person to make a product data request - a request to a DH to disclose CDR data that relates to products offered by the DH.
- The DH must disclose any required product data in response to the request. A DH cannot charge for required product data.
- The DH may disclose any voluntary product data in response to the request (but is not required to). A DH may charge for voluntary product data.
 - CDR Rule 1.12, Part 2 of the CDR Rules, Clause 3.1 of Schedule 3

Product data required to be disclosed

- Required *product data* must be disclosed in response to a product data request.
Required product data:
 - is within a class of information specified in the DI;
 - is about eligibility criteria, terms and conditions, price, availability, or performance of a product; and
 - for data about availability or performance - is publicly available; and
 - is product specific data about a product; and
 - is held in digital form.
- A DH must disclose required product data that is the subject of a request and that is otherwise contained on a DH's website or in the relevant disclosure document for the product.
- Clause 3.1 of Schedule 2 and CDR Rule 2.4.

How to report product reference data quality concerns

- If product reference data (PRD) is missing, inaccurate or incomplete, the data holder may be non-compliant
- PRD Non-compliance is a matter for the ACCC
 - **CDR Participants can report PRD compliance concerns to the ACCC via the CDR website: cdr.gov.au/contact-us > Make a report**
- Attach evidence to support your claims
- We will review in line with the CDR compliance & enforcement policy, available at: cdr.gov.au/resources

Next steps

—



Next steps and work to be done

—



We will learn as we go:

- First lets get started on addressing known issues (workshop Aug. 25)
- Develop Data Quality Assessment & Management Framework (ACCC / DSB / OAIC)
- Establish workshop cadence for issue discovery / discussion (DSB). Our processes for addressing data quality issues may be folded back into the normal maintenance cycles if/when the volume of data quality issues decrease

Thank you

—

<https://www.cdr.gov.au/>

<https://consumerdatastandards.gov.au/>

Appendix

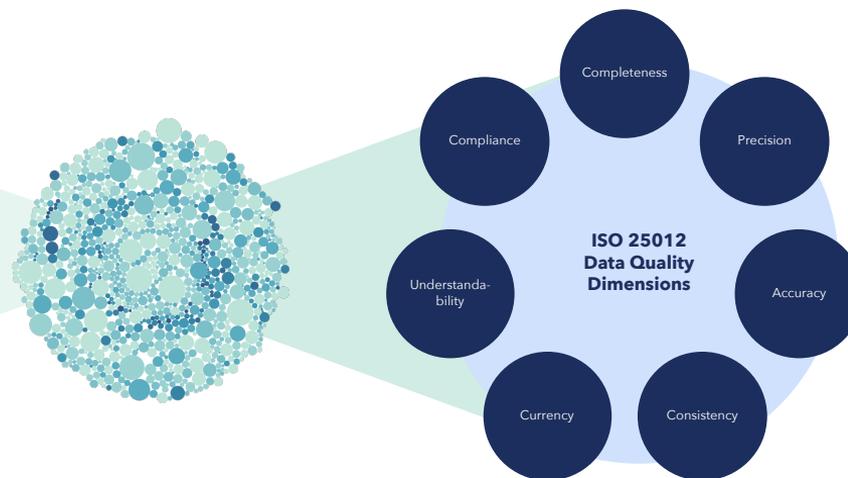
—

Defining data quality



ISO/IEC 25012 Data Quality Model defines data quality as:

- 1. Inherent Data Quality:** The degree to which quality characteristics of data have the intrinsic potential to satisfy stated and implied needs when data is used under specified conditions
- 2. System-Dependent Data Quality:** The degree to which data quality is reached and preserved within a computer system when data is used under specified conditions



Source: Subset of the ISO/IEC 25012 Data Quality Model

What problems are we trying to solve?

Non-compliance to provide required data

Incorrectly implementing the standards and rules

Ambiguous standards that reduce comprehension, compliance and consistency

Standards that offer too much flexibility which dilute consistency

Lack of association between the goods and services Data Holders offer and products the customer owns

Incorrect data held by Data Holders resulting in security/privacy issues or may incorrectly present products

Incomplete representation of products and consumer data

Developing a scalable process to sustainably manage data quality

Improving data quality = more innovation & better consumer outcomes

Goods and services can be *compared equally*

Consumers have *higher confidence* they are getting the value

Goods and services are comprehensively defined in a way that *adds value and drives innovation*

Consumers have *greater trust* in the CDR and applications in the ecosystem

Data Holders are encouraged and supported to provide data in the best interests of the consumer

Better interoperability across data recipients and data holders

Questions we need to consider



How do we:

- Make it easier to discover and report on issues?
- Address issues quickly so there isn't prolonged delay to innovation and competition?
- Achieve this without creating a change management nightmare?
- Encourage improvements in data quality in a timely manner?
- Use enforcement and penalties as a last resort: motivate the many to do good and make sure there is continual improvement without a heavy-handed approach?
- Strike the right balance between certainty and flexibility?
- Drive necessary changes to standards at a digestible and manageable pace?

What happens if we don't manage data quality?



- Higher implementation and integration costs
- Increased complexity and lower ROI leading to fewer ADRs and a smaller ecosystem
- Stymied innovation and agility
- Privacy breaches or financial loss to consumers
- Reduced value to consumers leading to limited uptake or sustained use
- Brand / reputation damage and reduced trust
- Constrained interoperability across sectors

A detailed look at the common issues

#	Issue Category	Data Quality Dimensions	Implications	Remedies
1.	Deliberate/ negligent non-compliance	<ul style="list-style-type: none"> • Accuracy • Completeness • Compliance • Currency 	DH fails to provide required data <ul style="list-style-type: none"> • Puts the onus on ADRs to scan PDSs against PRD data and • Not easy to discover all issues • Impact on consumers 	<ul style="list-style-type: none"> • Compliance/enforcement (ACCC)
2.	DH implementation error	<ul style="list-style-type: none"> • All dimensions 	Failure to correctly interpret the standards correctly <ul style="list-style-type: none"> • Analysis required to determine if the fault lies with the DH or with the standards. • Impairs one or many ADRs • Impact on consumers 	<ul style="list-style-type: none"> • Clarification of standards may be appropriate (see category 3) (DSB) • Escalation to category 1
3.	Ambiguous standards	<ul style="list-style-type: none"> • Completeness • Understandability • Consistency 	DHs populate the payloads in differing ways that are all at least arguably compliant or consistent with the standards <ul style="list-style-type: none"> • “Right” answer won’t always be easy to determine (particularly in the case of PRD and Account Detail payloads) • Reduced value to ADRs and consumers • Misleading comparisons • Increase in implied non-compliance 	<ul style="list-style-type: none"> • Creation of “data conventions” to be developed progressively with the community (DSB)
4.	Overly flexible standards	<ul style="list-style-type: none"> • Completeness • Understandability • Consistency 	Too much flexibility for DHs, resulting in data that is too variable and/or too poorly structured to be useful to ADRs <ul style="list-style-type: none"> • Increase in implied non-compliance 	<ul style="list-style-type: none"> • Clarification of the standards / rules (DSB / ACCC) • Changing the standards to be more prescriptive (DSB)
5.	Incomplete standards	<ul style="list-style-type: none"> • Completeness • Understandability 	<ul style="list-style-type: none"> • DHs may be consistent and complete but the standards / rules don’t sufficiently represent how products are compared • Reduced value to ADRs and consumers • Misleading comparisons • Consumers don’t have access to all of their data 	<ul style="list-style-type: none"> • Clarification of the standards / rules (DSB/ ACCC) • Changing the standards / rules to be more prescriptive (DSB / ACCC) • Data Conventions (DSB)
6.	Lack of product relationship	<ul style="list-style-type: none"> • Understandability • Currency • Precision 	Lack of linkage / association between products DHs offer in market versus the products customers own (i.e. accounts). May result from grandfathered products, inconsistent product naming or negotiated product features/rates when a customer purchases a product.	<ul style="list-style-type: none"> • Less easy to resolve • Perception of data quality issues
7.	Incorrect data held by DHs	<ul style="list-style-type: none"> • Accuracy • Precision 	DH may present data correctly but the data itself is incorrect – for example often times joint accounts may have one spouse’s details incorrectly recorded against the other spouse.	<ul style="list-style-type: none"> • Compliance/enforcement (ACCC/ OAIC) • Notify ADRs of data changes

Privacy Safeguard 11: Quality of CDR data



Privacy Safeguard 11 requires entities to:

“take reasonable steps to ensure that the CDR data is, having regard to the purpose for which it is held, accurate, up to date and complete”

- Privacy Safeguard 11, together with CDR Rule 7.10, also requires a data holder to inform a consumer in the event that incorrect CDR data is disclosed, and disclose corrected CDR data to an ADR (where requested by a consumer).
- For further information, see [Chapter 11](#) of the OAIC’s CDR Privacy Safeguard Guidelines.