d-fine

Mastering Data Lineage under RDARR

Leveraging It in Business Use Cases to Foster a Data Driven Organisation

analytical. quantitative. tech.

2026

Content

Mastering Data Lineage under RDARR, 2026

01.

Data lineage: From a regulatory topic to a business enabler

In this whitepaper, we discuss the benefits, regulatory requirements and different approaches to data lineage documentation. Data lineage is currently widely debated in the market and has become a key topic for banks seeking to strengthen data management, transparency and alignment with the expectations set out in the ECB's Guide on effective risk data aggregation and risk reporting (RDARR Guide)[1]. The main challenge is to find an approach that balances regulatory expectations and implementation effort with tangible business value while strengthening the organization's AI readiness.

Our view

Three design principles for effective data lineage that maximize business value and keep implementation efforts proportionate:

1 Business relevance is key – Lineage should be business-led and prioritise data objects that are business-critical.
2 Flexible granularity enables fit-for-purpose approaches – The level of detail should be based on a clear materiality and criticality assessment.
3 Targeted technical drill-downs bridge the gap – Where needed, the business-view should be complemented by drill-downs in the technical layer.

Additional success factors:

- The interpretation of "front" in "front-to-end" should be clearly defined, justified and applied consistently.
- Sustainable lineage requires clear ownership and integration into change processes.
- Data lineage creates value beyond compliance: better data quality, faster change impact assessment and AI readiness and a more data-driven organisation.

Transparency over data flows – the data's origin and how it is transformed and aggregated across systems – is one of the key enablers of business value for banks of all sizes, although the implementation approach should reflect the institution's complexity, risk profile and regulatory exposure. In day-to-day operations, data lineage improves data quality by enabling the implementation of targeted data quality checks and faster root-cause analysis, thereby reducing overall remediation efforts by up to 30%[2]. By providing a shared view of data flows across business and IT, it reduces manual investigation and reconciliation efforts and strengthens operational resilience. Data lineage also enhances change capabilities.

1 Please see RDARR-Guide.

2 Numbers are estimated based on internal project experience of different change and run projects in various banks.

A transparent view of how data moves across systems makes the overall architecture easier to manage and allows for instance changes to systems, data fields or transformation logic to be assessed more efficiently and reduces the effort needed for time-consuming data discovery activities. This helps identify critical dependencies and downstream impacts early, reducing implementation risks and efforts in transformation projects. In our experience, available data lineage can improve efficiency of implementation projects by up to 20%[2] by reducing time-consuming data discovery efforts. Data lineage also supports regulatory compliance with BCBS 239 and local regulations, as well as alignment with the expectations outlined in the ECB's RDARR Guide.

As illustrated in Figure 1, effective lineage requires combining business and technical perspectives: the business view[3] explains how data is used and transformed conceptually, whereas the technical view shows the physical data flow.[4] Both represent horizontal lineage, showing how data moves across systems from front to end. The business and technical lineage are connected via the vertical lineage, which maps business objects to their technical representations and vice versa. Business-led data lineage, with flexible granularity based on criticality and materiality and targeted technical drill-down capabilities, maximizes usability and business value while keeping implementation efforts proportionate.

This whitepaper focuses on the expectations regarding lineage documentation for data objects within the scope of application[5] as defined by the RDARR Guide. In particular, we discuss expectations around data-attribute-level documentation and front-to-end lineage, as these raise practical questions regarding granularity, maintainability and proportionality. In general, we recommend extending data management activities beyond the regulatory scope, as well-managed data creates tangible value beyond compliance.

3 The logical data flow, herein referred to as the business lineage.

4 While the documentation of the technical lineage is primarily driven by regulatory expectations and is required for the effective implementation of data quality checks, business lineage is the key to enabling concrete use cases.

5 Please refer to Shaping the RDARR Scope of Application.

02. At a glance: From requirements to practice

Since the publication of the RDARR Guide in May 2024, its expectations have triggered extensive implementation and remediation activities across the market. While supervisory priorities may evolve[6], robust data management will remain a lasting business requirement.

The ECB expects the lineage to be documented

1 at the **data-attribute level**, where *"[...] data attributes are often stored as a column in a table and are used in the technical mapping [...]"*[7] and
2 in a **front-to-end** manner, i.e. the documentation shall start from *"[...] data capture [...] including extraction, transformation and loading"*[7].

These expectations are intended to enable banks to understand the impact of data quality issues arising along the entire processing chain on the Key Risk Indicators (KRIs) and Critical Data Elements (CDEs), thereby supporting sound decision making and effective steering. However, as of now, banks typically do not have a granular data-attribute lineage along the entire processing chain, front-to-end.[8]

Solutions observed in the market usually aim for an institution-specific sweet-spot between the following two extremes:

1 Capturing of the (technical) lineage in an automated manner in exceptional cases.[9]
2 Mostly manual lineage documentation, usually starting from a business view.[10]

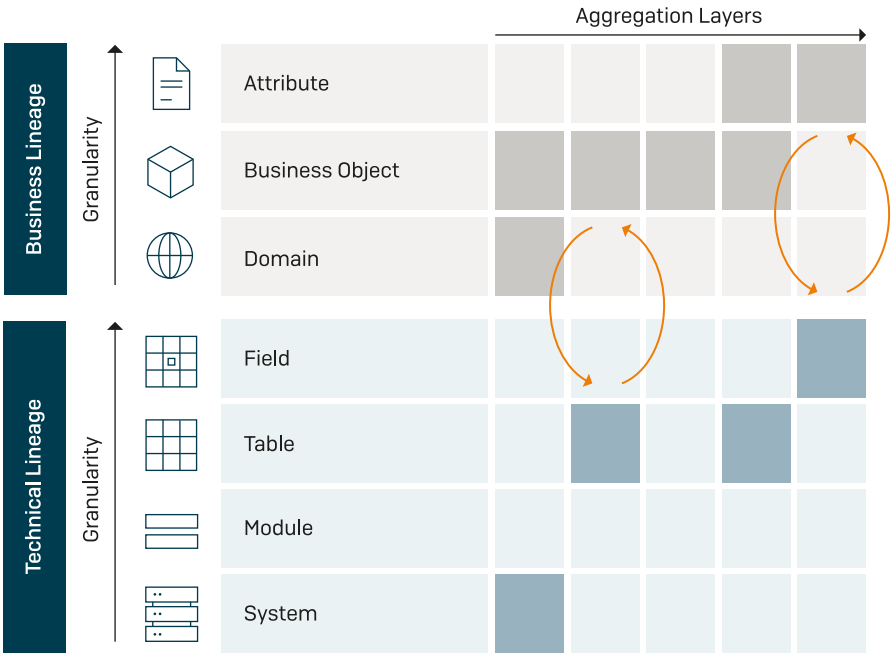

Figure 1: Sketch of a documented lineage. The granularity of the Business Lineage is adjusted based on the criticality and is complemented by the technical lineage.

AI-based approaches are increasingly explored to derive glossaries, mappings and lineage relations from existing documentation. While promising, their effectiveness depends heavily on documentation quality and requires validation by subject-matter experts.

6 Please see supervisory priorities 2025-2027 and 2026-2028.
7 Please see section 3.4 article 3 in the RDARR-Guide. Note that the term "end-to-end" is also synonymously used in the market.
8 Front as in data capture or sourcing.
9 Those approaches typically work in a homogenous system landscape with few technical breaks in the data flow and does stand-alone not provide meaningful business context.
10 While exhibiting more manual workload, this approach provides most business insight and captures relevant information by subject matter experts.

03.

Balancing the granularity: Regulatory requirements, implementation efforts and usability

Meeting the ECB's expectations can require significant implementation and maintenance effort, given system breaks, third-party software, legacy architecture and the sheer number of data elements[11]. The key challenge is to document lineage in sufficient detail while avoiding overengineered solutions that are difficult to maintain and provide limited additional value.

Further details on the required granularity:

In practice, many banks document the lineage at the business object level and introduce a hierarchy with different levels of granularity (e.g., Street as part of Address, Address as part of Client) – the ability to adjust the granularity as required supports the development of fit-for-purpose solutions. Data lineage may be documented at a more detailed level close to the reports and models in scope, while upstream data flows may be captured at a more aggregated level. Granularity is then increased where necessary at critical points along the data flow. Others are focusing their lineage documentation directly on the (level of) CDEs. Both strategies provide the business lineage and are in general consistent, as the CDEs are typically determined by subject matter experts using a risk- or materiality-based approach.

As this approach primarily provides the business lineage it should be complemented by a sufficient number of entry points via the vertical lineage into the technical lineage, to allow targeted drill-downs where necessary and enable the implementation of DQ-checks. Following that logic, many banks also distinguish between the business view and the technical view on CDEs.

11 To get a feeling regarding the complexity and required efforts let's look at the highly simplified example: Assume a layered processing chain consisting of 5 aggregation steps. In each aggregation step the data elements are derived using 3 data elements from the preceding layer. Starting from the KRI which is derived using 3 data elements in the preceding layer, which in turn are derived using 3 data elements each, we finally end up with 3^5 = 243 data elements in the last layer (front office). Of course this is a highly simplified example, as on the one hand lineage of different KRIs tends to overlap and the number of data elements in the front-office systems is of course limited.

04.

Where does the lineage commence: Defining the front in front-to-end

Extending the lineage documentation to the front office systems also creates significant practical challenges for banks. For instance, the system landscape often becomes increasingly heterogeneous and fragmented toward the front office, the number of owners and stakeholders increases, and the higher frequency of system changes makes it difficult to keep documentation accurate and up to date – circumstances that can substantially increase the required documentation effort.

Further details on front-to-end:

While supervisors conceptually understand "front" as the point of data origination, many institutions consider the systems where data is stored persistently for the first time as the "front". Hence the documentation of the data lineage often commences from the system of record (SoR). It should be noted that, from the point of first central persistent data storage, the range of feasible DQ-checks increases significantly (i.e. completeness or consistency checks, temporal consistency). Banks should define "front" in their data management policy, justify the interpretation and ensure consistent application across the organisation.

When it comes to the identification of CDEs, banks often introduce a hierarchy of CDEs along the processing chain and limit the number of levels in this hierarchy in order to keep both the number of CDEs and the overall complexity manageable. From a regulatory perspective, supervisory communications point towards an expectation that, where applicable, CDEs are identified in a consistent manner along the entire processing chain. CDEs should not be introduced at arbitrary points in the lineage without a clearly traceable source in the upstream data capture and front-office systems.

05.

Data lineage and the operating model

While concepts and frameworks outlining the requirements for the lineage documentation are typically owned centrally by the data management function, the operationalisation / roll-out are often performed decentralised / federated in data mesh-like setups. Usually, the data owners are accountable for the lineage documentation, its completeness and maintenance, while data stewards (or similar roles) are responsible on an operational level.[12]

12 Often data owners act primarily from a business perspective and data stewards on a technical level – of course this depends on the specific organizational setup.

In practice, lineage documentation is initiated from in-scope reports, models and KRIs, and then traced upstream along the data flow. Data ownership may change at material aggregation or transformation steps in the processing chain, particularly where responsibilities shift between business domains or functional areas.[13]

Lineage documentation should not be treated as a one-off remediation exercise, but embedded into change, release and data quality management processes. A consistent level of implementation is supported by centrally implemented quality assurance steps – for instance by automated controls on metadata (computational data governance), awareness measures or sample-based deep dives of the documented lineage.

06. Effective data lineage: Usability, transparency and compliance

A fit-for-purpose lineage approach should be driven by business needs and based on criticality and materiality. Extending data management beyond the regulatory minimum strengthens AI readiness, supports change initiatives and improves run-the-bank processes. Robust data lineage therefore supports better decision-making[14] and fosters a data-driven organisation.

Designing fit-for-purpose solutions requires a sound understanding of regulatory expectations, market practices and aggregation processes. Efficient implementation, in turn, depends on a thorough understanding of the logical data flow and its underlying technical implementation. The most effective lineage approaches are business-led, proportionate and embedded in day-to-day governance and change processes. They meet supervisory expectations while creating tangible value for operations, change and analytics. d-fine supports banks in translating these principles into fit-for-purpose lineage frameworks, operating models and implementation roadmaps.

13 Without integration into change and release processes, lineage documentation quickly becomes outdated.

14 For instance, if the probability of default of a new potential client is estimated inaccurately due to poor data quality, the bank takes either unnecessary risks or misses out on potential business.

Authors

Dr Philipp Altmeyer
Manager & Expert in Data Management
d-fine GmbH, Frankfurt
Philipp.Altmeyer@d-fine.com

Maik Schock
Senior Manager & Expert in Data Management
d-fine GmbH, Dusseldorf
Maik.Schock@d-fine.com

Dr Sascha Hügle
Partner & Expert in Data Management
d-fine GmbH, Frankfurt
Sascha.Huegle@d-fine.com

Dr Ingolf Segger
Senior Manager & Expert in Data Management
d-fine GmbH, Frankfurt
Ingolf.Segger@d-fine.com

d-fine

Berlin
d-fine GmbH
Kurfürstendamm 21
10719 Berlin
Germany
berlin@d-fine.de

Dusseldorf
d-fine GmbH
Gustaf-Gründgens-Platz 5
40211 Dusseldorf
Germany
duesseldorf@d-fine.de

Frankfurt
d-fine GmbH
An der Hauptwache 7
60313 Frankfurt
Germany
frankfurt@d-fine.de

Hamburg
d-fine GmbH
Am Sandtorpark 6
20457 Hamburg
Germany
hamburg@d-fine.de

London
d-fine Ltd
14 Aldermanbury Square
London, EC2V 7HR
United Kingdom
london@d-fine.co.uk

Milan
d-fine s.r.l.
Piazza Sigmund Freud, 1
20124 Milan
Italy
milano@d-fine.com

Munich
d-fine GmbH
Friedenstraße 22b
81671 Munich
Germany
muenchen@d-fine.de

Stockholm
d-fine AB
Kungsgatan 9
11143 Stockholm
Sweden
stockholm@d-fine.se

Stuttgart
d-fine GmbH
Kriegsbergstraße 11
70174 Stuttgart
Germany
stuttgart@d-fine.de

Utrecht
d-fine B.V.
Stadsplateau 7
3521 AZ Utrecht
Netherlands
utrecht@d-fine.nl

Vienna
d-fine Austria GmbH
Seilerstätte 13
1010 Vienna
Austria
wien@d-fine.at

Zurich
d-fine AG
Brandschenkestrasse 150
8002 Zurich
Switzerland
zuerich@d-fine.ch