

2.4.3 Observations with Provenance

[Return to 2.4 Existing Functional Capabilities in Reporting Systems](#)

Reporting systems increasingly include a functional capability that associates extracted values or observations with provenance information. This provenance links each observation to specific source artefacts, document locations, fields, or extraction contexts and records how, where, and under which conditions the system derived the observed fact.

At the architectural level, provenance is more than ancillary metadata. It provides the mechanism that connects observed facts to their evidentiary origins and enables traceability from interpreted results, through extraction processes, back to the original reported artefacts.

Provenance information may include:

- Source document identifiers
- Structural or field locations
- Extraction methods
- Processing timestamps
- Confidence or uncertainty measures
- Tool, model, or component identifiers
- Processing context
- Version information

Many existing systems capture provenance inconsistently or treat it as an implementation detail of extraction tooling. They may store provenance separately from [semantic](#) representations and fail to link it systematically to interpretation, validation, or analysis logic.

As a result, downstream systems may consume observed facts without enough context to:

* Assess their reliability * Reconstruct how the system produced them * Reinterpret them under revised semantic definitions * Compare observations generated by different extraction methods * Explain their contribution to derived conclusions * Distinguish extraction uncertainty from interpretive uncertainty

From an architectural perspective, this functional component ensures that every observed fact remains explicitly traceable to its source and extraction context throughout the reporting lifecycle, including when observations cross system, organisational, or jurisdictional boundaries.

This responsibility includes:

- Maintaining stable identifiers
- Preserving links to source artefacts and source locations
- Retaining extraction context and processing history
- Preserving linkage across semantic versions
- Supporting reevaluation under revised interpretive contexts
- Maintaining evidentiary continuity as observations move between components and systems

By treating observations with provenance as a distinct functional capability, the [Reference Architecture](#) establishes traceability as a governed architectural property rather than as an afterthought or tool-specific feature.

This separation supports auditability, reproducibility, controlled semantic evolution, and responsible use of artificial intelligence-assisted extraction techniques, particularly in environments where interpretations may change while the underlying reported artefacts remain fixed [R10], [R11].

© 2026 Dido Solutions, Inc. and Jackrabbit Consulting, Inc.

From:
<https://wiki.didosolutions.com/> - Dido Solutions, Inc Wiki

Permanent link:
<https://wiki.didosolutions.com/dido/01-fdis-ra/02-background-and-motivation/02-4-existing-functional-capabilities-in-reporting-systems/02-4-2-information-extraction-and-observation/start?rev=1783881780>

Last update: 2026/07/12 11:43

