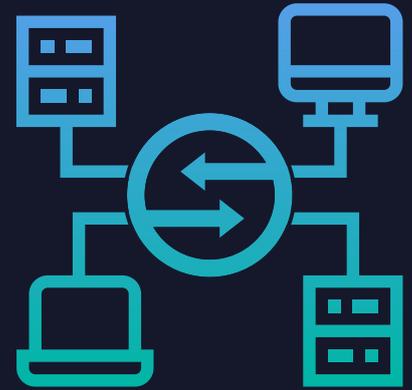# Solidatus

Importing into Solidatus

Solidatus is an award-winning web-based data lineage application which enables teams to collaboratively map and analyse their data landscape at scale, entirely within the browser. Solidatus' highly tuned user interface speeds up the manual analysis and documentation often required on data lineage projects, especially for future-state planning and what-if analysis.

Defining lineage and associated metadata in Solidatus is easy and can be done manually, by importing from one of the supported sources, or by using a custom connector. Particularly when modelling an existing data flow, efficiencies can be gained by using existing documentation or metadata, for example in Excel, which can be loaded into Solidatus to both bootstrap a data lineage model and bring more understanding to an otherwise opaque and complex format for maintaining lineage. Additionally, there may be introspection points available, such as databases, governance applications, ETL systems and internal tooling from which schemas can be extracted for loading into Solidatus.

CONNECTIVITY

- Solidatus supports importing structure, mappings and annotations from many data sources including Excel, relational databases, clouds (e.g. GCP's BigQuery and DataFlow), ETL tools, graph databases, RDF, CSV, JSON, XML, batch processors (e.g. Autosys and Control-M) and others.

- Solidatus can extract schemas from all mainstream relational databases and has extensive support for importing both structure and lineage from other applications, SQL stored procedures and other data sources through our extensive partner network.

- Import metadata from graph databases such as Stardog and Virtuoso using the SPARQL connector or RDF importer. Combine this with collaboratively documented lineage for powerful end-to-end lineage across many technologies and manually mapped processes.

- Use the flexible Excel/CSV importer with one of the templates to import existing documentation or use Excel as a stepping-stone format for importing from other sources.

FLEXIBILITY

- Any component in a Solidatus model can be imported, either separately or in combination with other parts of the model. For example, a spreadsheet can be used to import a single database, a set of mappings, just some properties, or all of the above.

- The Solidatus API can be used to automate data lineage documentation by programmatically submitting data lineage metadata to the JSON-based HTTP endpoints. API changes are versioned, audited and can be augmented with manually maintained metadata and lineage.

- When importing entities, mappings and properties there are several in-app customisations which can be made. For example, where the imported entities already exist, the importer can merge/update the two versions, overwrite the existing version, or duplicate the entities.

- The same dataset can be imported at regular intervals (or on an event) and Solidatus will detect and highlight the change in the lineage visualiser. It will also version the metadata in Solidatus' version control system. This can be used to perform impact analysis of database schema changes, for example.

ENTERPRISE READY

- Mappings and metadata spread across siloed data stores and ETL tools can be loaded into a single Solidatus model, resulting in true end-to-end attribute-level data lineage which is easy for business users to interface with and contribute to.

- Automatically ingested lineage harvested from modern applications can be dovetailed into manually maintained lineage for legacy processes allowing progression over time to fully automated lineage.

KEY POINTS

▶ Support for a wide variety of data sources

▶ Quickly and easily import lineage using Excel templates

▶ Extract schemas from relational databases and the cloud

▶ Detect and visualise change (e.g. schema upgrades)

▶ Integrate Solidatus into your existing governance ecosystem