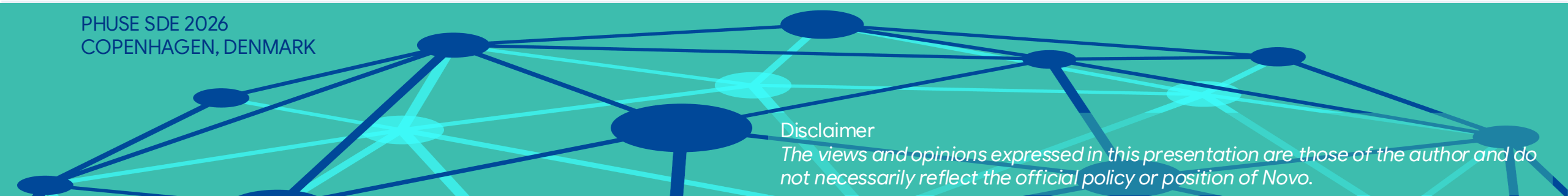
Stacking with Ease in a Cloud-Based Data Lakehouse

Vicky Poulsen, Standards Development Director

Novo Nordisk A/S

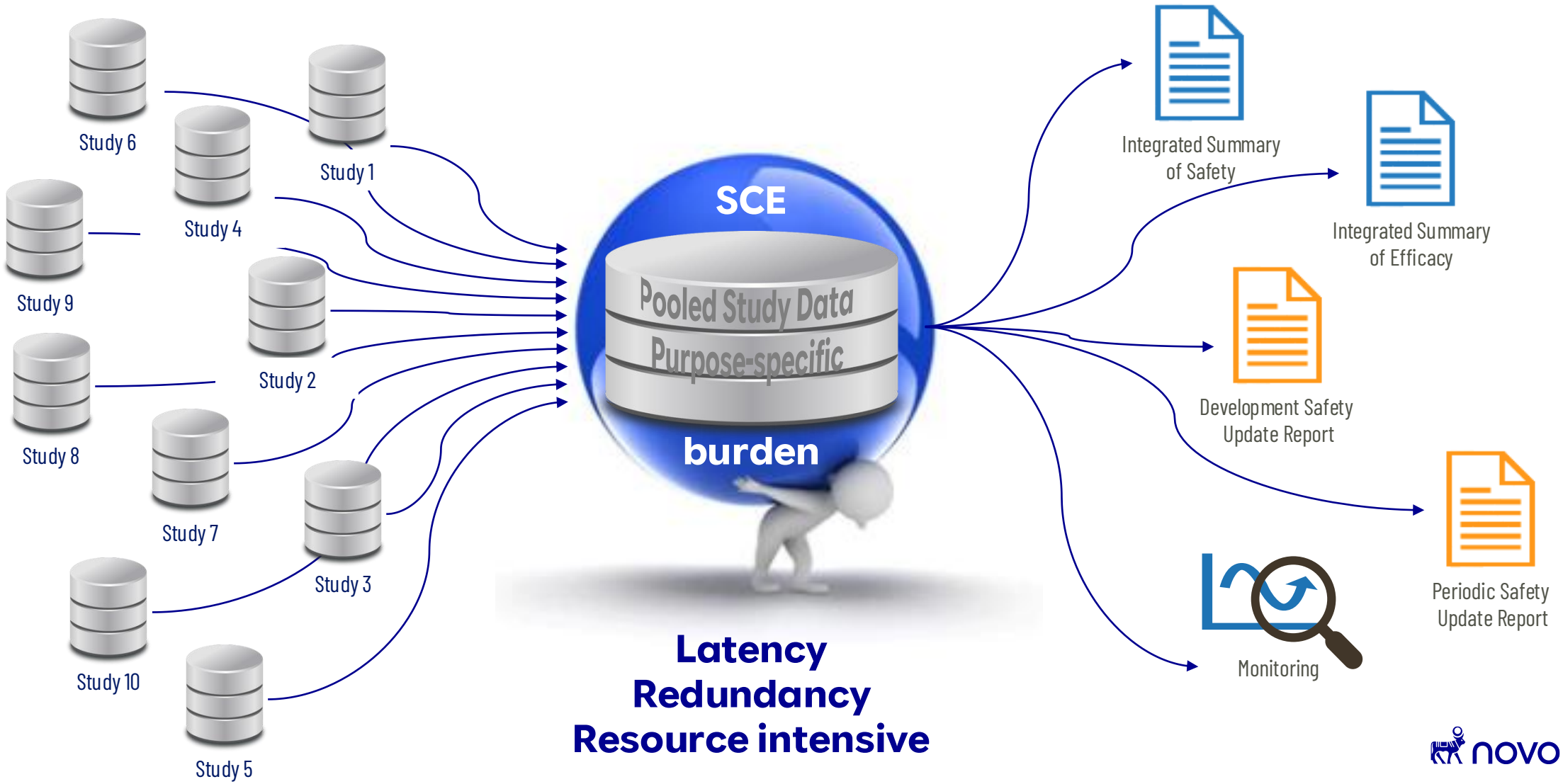


PHUSE SDE 2026
COPENHAGEN, DENMARK

Disclaimer

The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of Novo.

Data Pooling for Integrated Reporting/Monitoring



Modernising Clinical Data Flow

Papillons SDTM Generation Framework | SAS Viya | Databricks | OneLake

Lakehouse Architecture
Unified data and orchestration platform

-enabled Integration
Seamless automation

DuckDB Extensions
In-process on-storage file access

Spark Engine
Reading and writing to delta tables

AI Assistant
End-user support



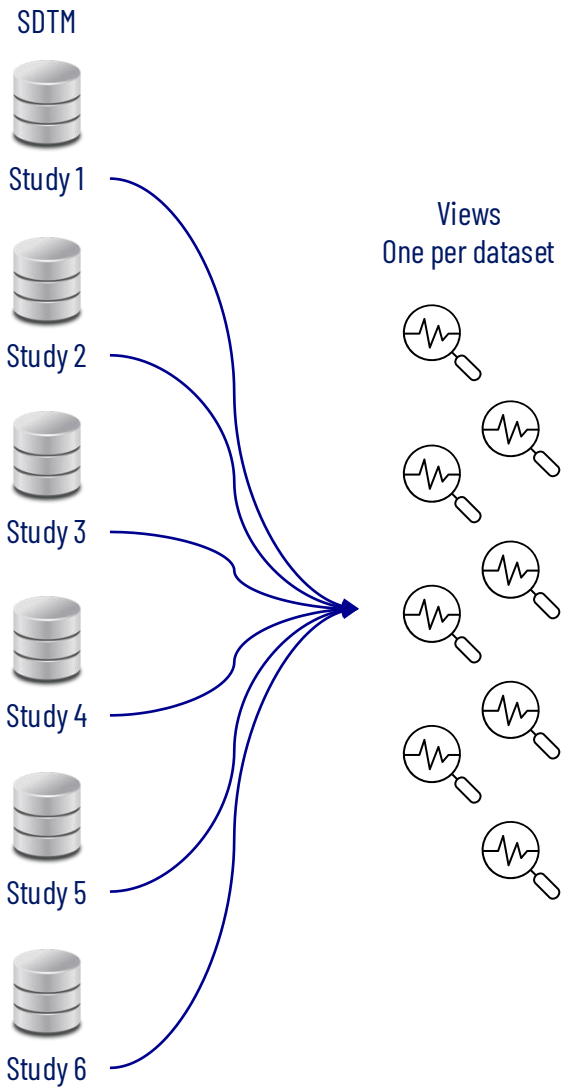
UNIX SCE
hp UX
sas 9.4

Papillons

2027

novo

SDTM Stack | A Data Product for Data Pooling



View is a virtual table defined by SQL query

What
- Study Metadata
- SDTM

How
- View

Where
- Databricks

Benefit
- No data duplication
- Synchronised with source data
- Support language agnostic use
- Out-of-the-box feature
- GxP compliant
- Efficient, governed process

Prerequisite
- Data in Study Catalog
- Stable connection
- Scalable compute

Concerns
- Performance
- Access
- Stability

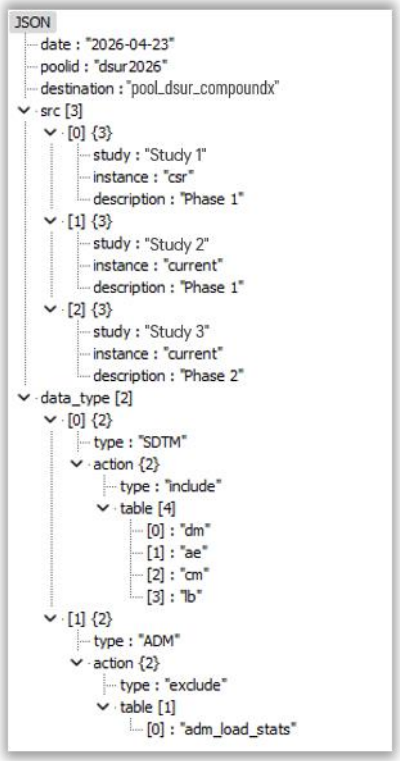
Potential
- ADaM Pool
- Continuous monitoring
- Integrated reporting
- ad hoc analysis

SDTM Stack | An End-to-End Solution

Request | Process | Deploy | Finalise

Stack Definition

Provided by requestor



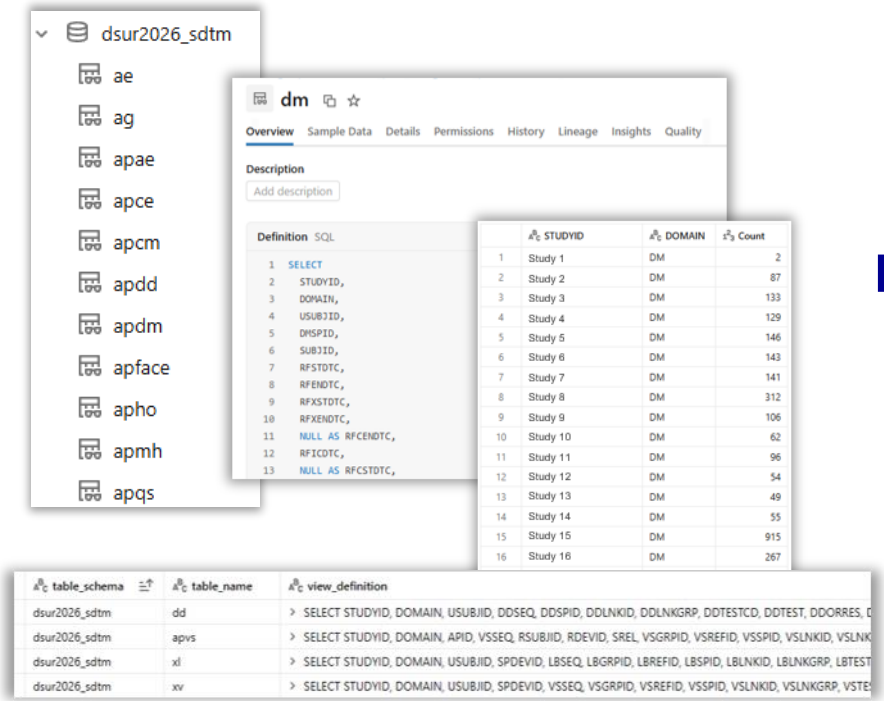
Stack Engine

Developed and maintained as a Papillons component

Triggered by Stack Definition

Stack Views

*Populated in **confidential** POOL catalog in Databricks*

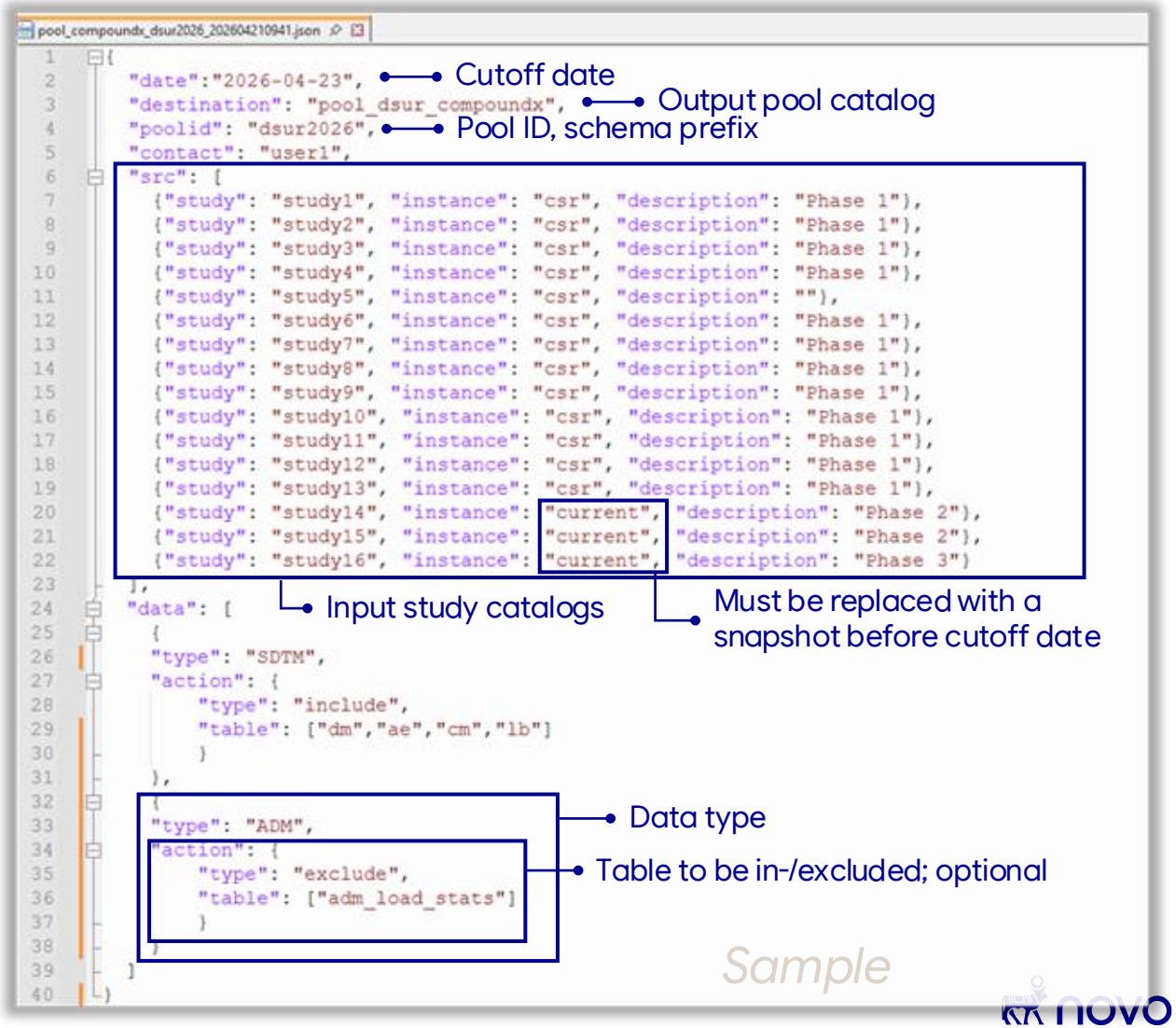


Consumer

*With access to source and **confidential** POOL catalog*

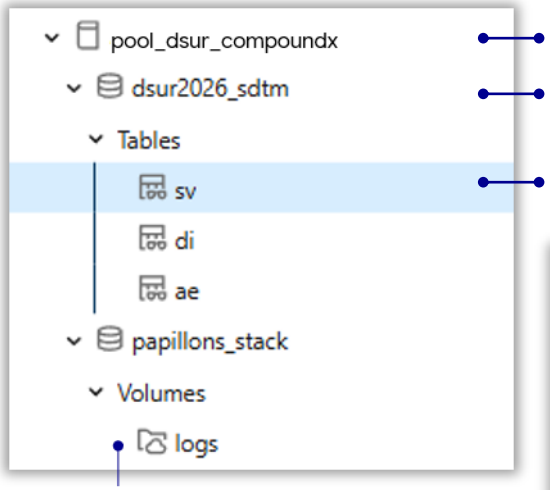
Stack Definition | Self-documenting

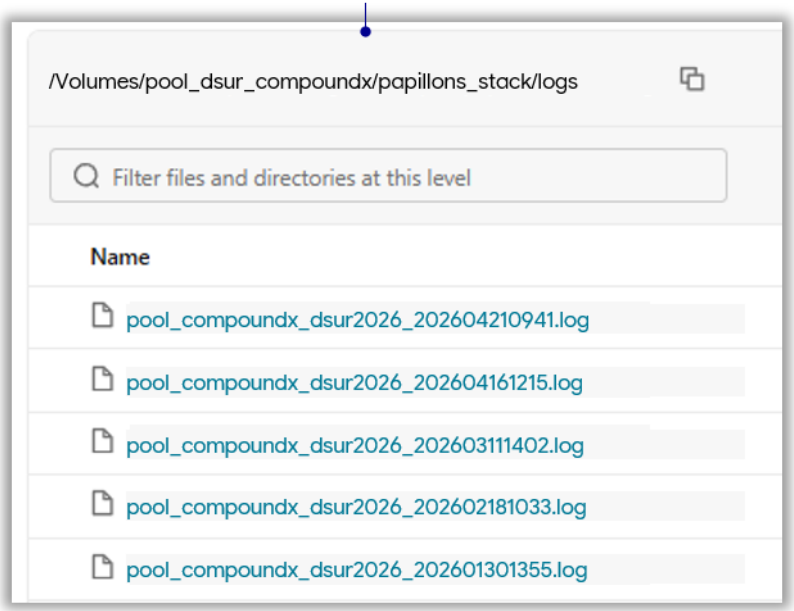
- One definition per deliverable specifying studies and data to be stacked
- JSON: a very simple structure, template available
- Source to SDTM Stack **must** be a snapshot on the specified cutoff date

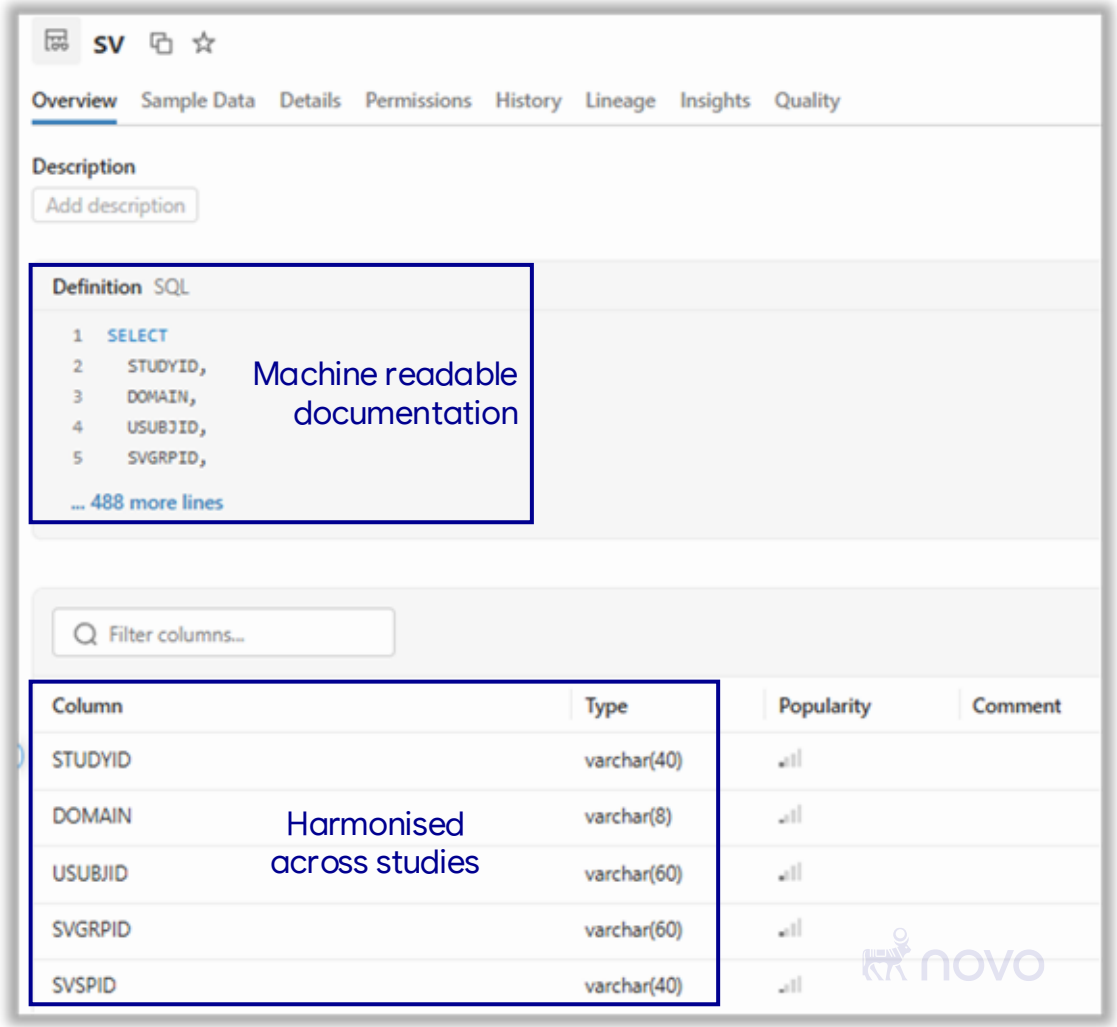


Stack Views

- Refresh when Stack Definition is updated
- Locked upon request







Learnings and Reflections

- SDTM Stack maximises on new technologies
  - Asynchronous workflow
  - No data latency
  - No data redundancy
  - Data lineage – clear traceability
  - Flexible and scalable
- Challenge the conventional way of working
  - Self-service, self-documenting approach
- Avoid technology bias
  - Choose the right tool for the job
- Stability remains the biggest challenge in asynchronous workflow
  - Invalid views when source data is missing due to data upload failure



Thank you!

Vicky Poulsen
(vicp@novonordisk.com)

Credit:
Standard Programming Team, SSI