Netwerkdag Statistiek Vlaanderen – NBB SDMX & VTL

Frederik Van Hecke

21 Maart 2024

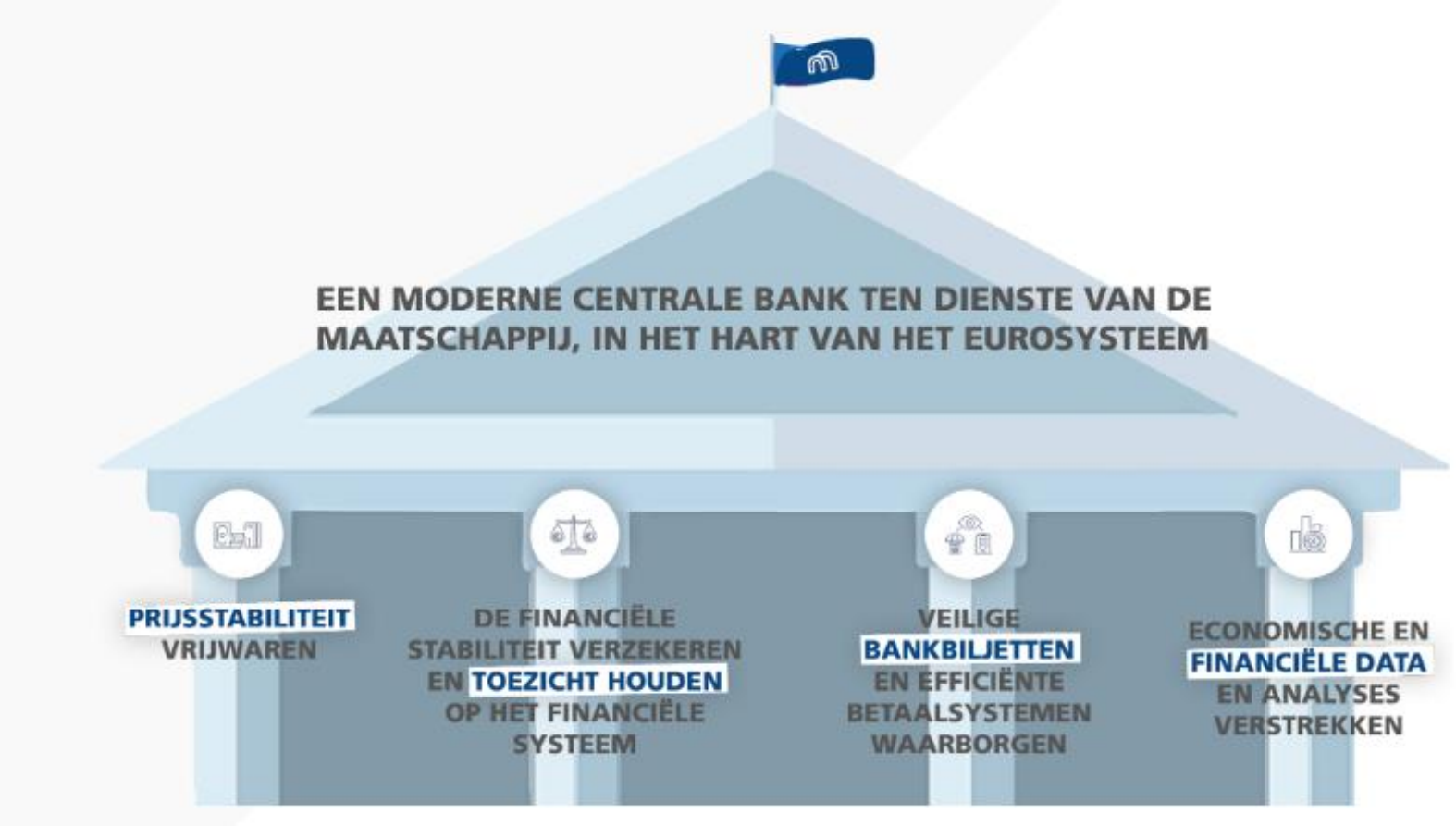
National Bank OF BELGIUM
Eurosystem

Agenda

Hoe zetten we verdere stappen vooruit in het toepassen van kwaliteitsstandaarden op de productie en publicatie van openbare statistieken?

Wat zijn hier de uitdagingen/valkuilen/opportuniteiten voor de individuele entiteiten en voor het netwerk als geheel?

Introduction



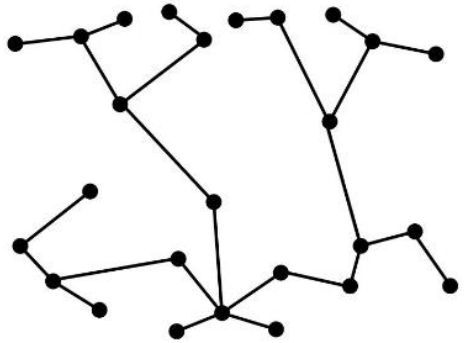
- Studies
- Statistieken
- Documentatiebestand
- Balanscentrale
- Centrale voor kredieten aan ondernemingen
- Centrale voor kredieten aan particulieren

Agenda

Hoe zetten we verdere stappen vooruit in het toepassen van kwaliteitsstandaarden op de productie en publicatie van openbare statistieken?

Wat zijn hier de uitdagingen/valkuilen/opportuniteiten voor de individuele entiteiten en voor het netwerk als geheel?

Quality Control – Current state



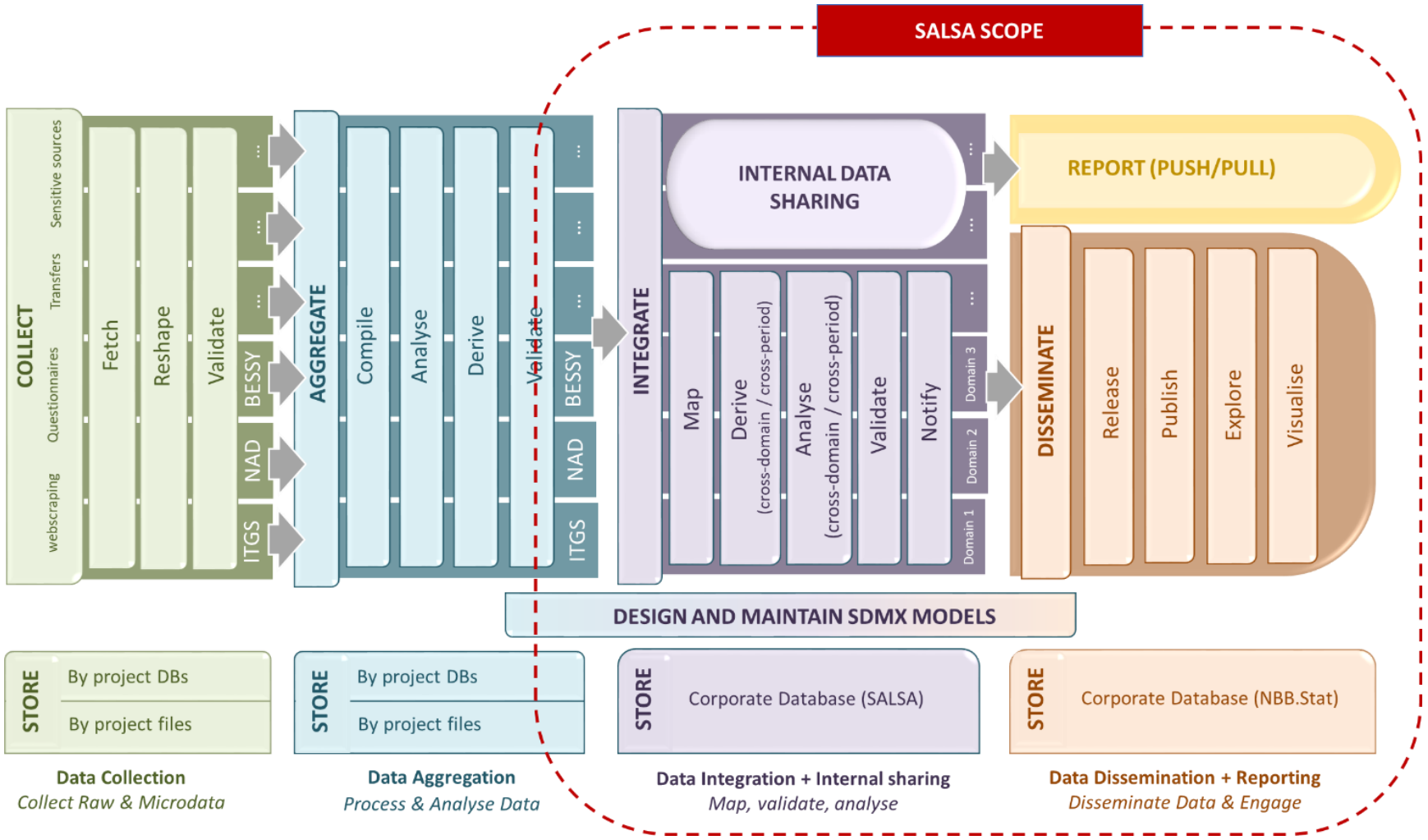
DECENTRALIZED

Need for modernization

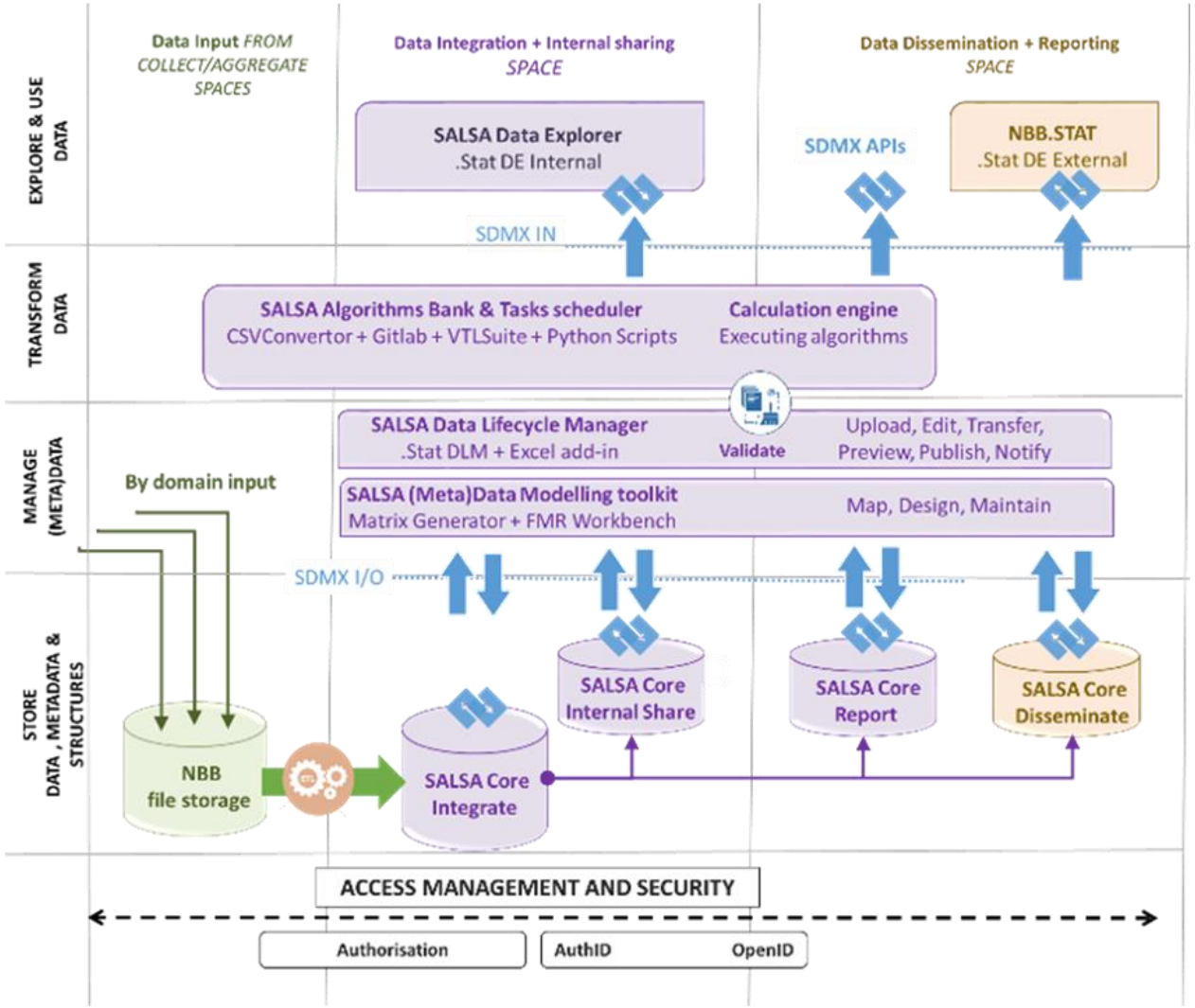
Business Case

- 2020: legacy application needs to be replaced
- Requirements:
  - **Process improvement**
  - **Single point of information** for aggregated statistics:
    - To be published/disseminated
    - International Organisations: automatic publication/dissemination
    - Internal (and/or intermediary) data for NBB data owners
  - **Data management/remodelling**
    - Series/observations
    - Datasets (data + metadata)
  - **SDMX based**
  - **Improved performance**
  - **Intuitive**
  - **More efficient software maintenance**
  - **Connectivity with other applications**
  - **Improved validation & quality of data**

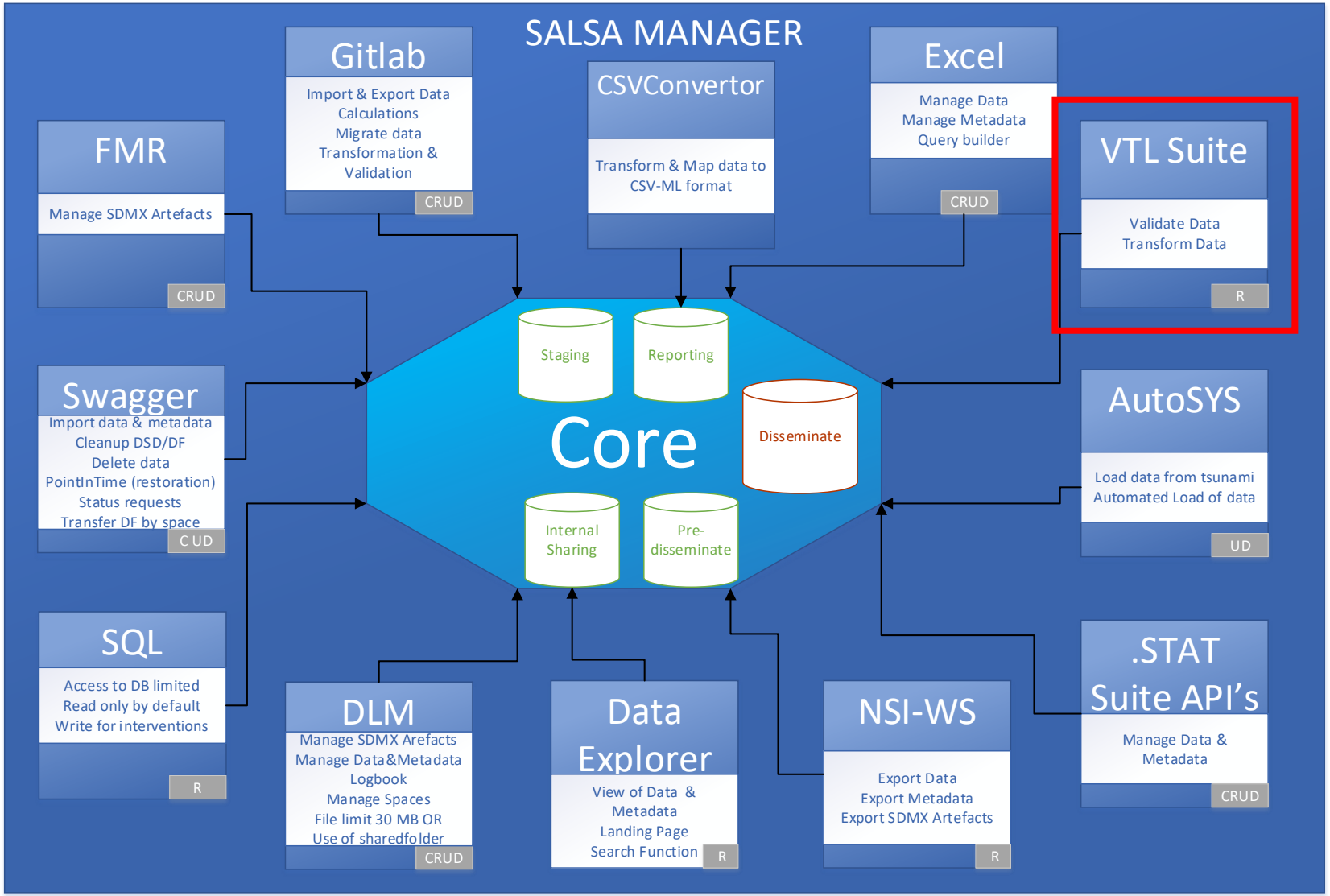
SALSA project scope



High level infrastructure



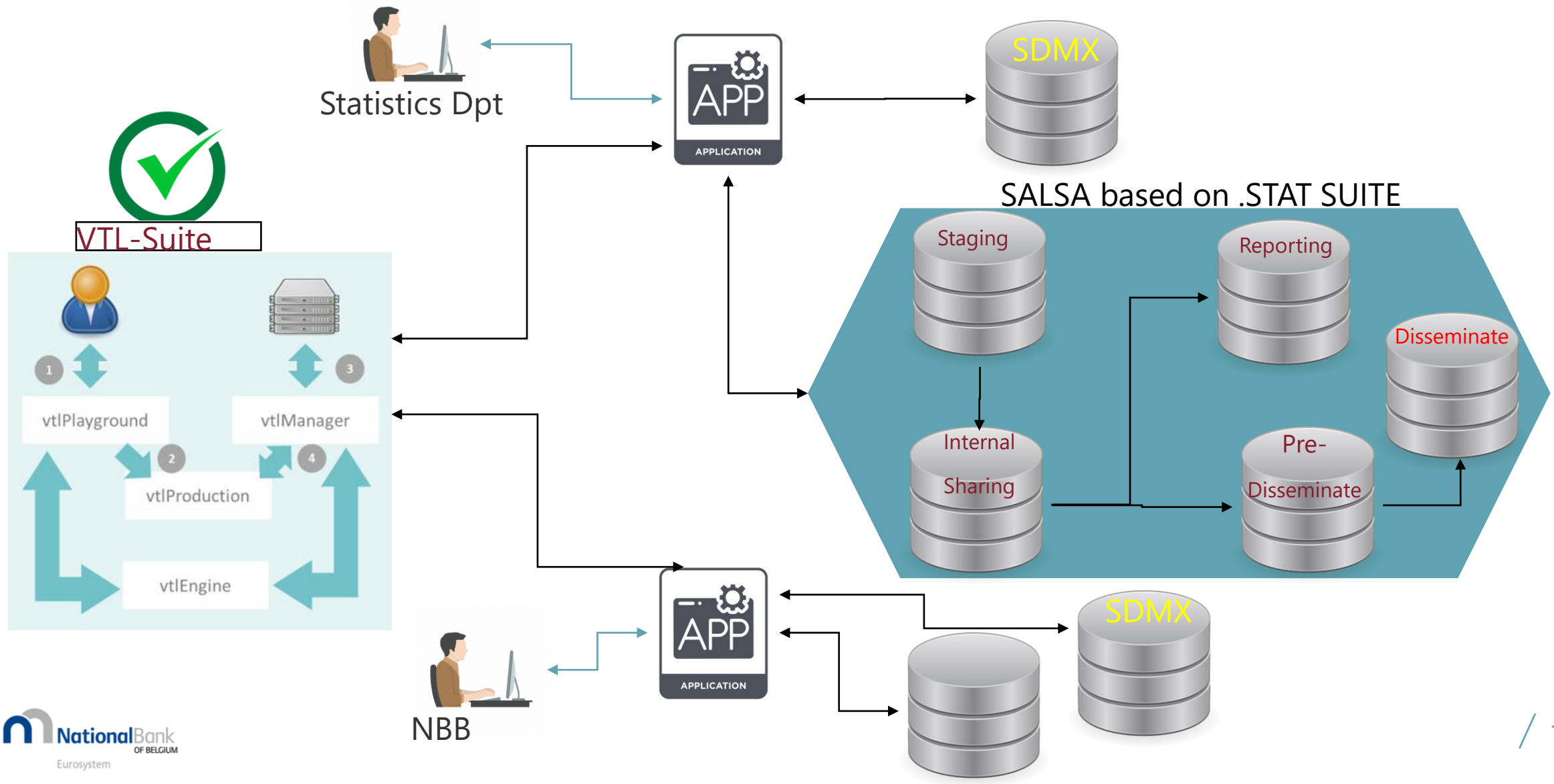
SALSA Tools


SALSA MANAGER
Gitlab
Import & Export Data
Calculations
Migrate data
Transformation & Validation
CRUD
CSVConvertor
Transform & Map data to CSV-ML format
Excel
Manage Data
Manage Metadata
Query builder
CRUD
FMR
Manage SDMX Artefacts
CRUD
VTL Suite
Validate Data
Transform Data
R
Staging
Reporting
Core
Disseminate
Internal Sharing
Pre-disseminate
Swagger
Import data & metadata
Cleanup DSD/DF
Delete data
PointInTime (restoration)
Status requests
Transfer DF by space
C UD
AutoSYS
Load data from tsunami
Automated Load of data
UD
SQL
Access to DB limited
Read only by default
Write for interventions
R
DLM
Manage SDMX Arefacts
Manage Data&Metadata
Logbook
Manage Spaces
File limit 30 MB OR
Use of sharedfolder
CRUD
Data Explorer
View of Data & Metadata
Landing Page
Search Function
R
NSI-WS
Export Data
Export Metadata
Export SDMX Artefacts
R
.STAT Suite API's
Manage Data & Metadata
CRUD

Why SDMX & VTL

- Already using SDMX for Reporting
- International Standard (ISO) for exchanging statistical data
- Large Community behind SDMX
- Use of .STAT Suite developed and maintained by OECD
  - NBB.STAT data portal already based on .STAT Suite
- SDMX offers a quality standard through the information model
- Structural validation through SDMX
- Content validation through VTL (Validation & Transformation Language)
- User friendliness
- Re-usability

Full NBB Solution for Validation & Transformation of Data

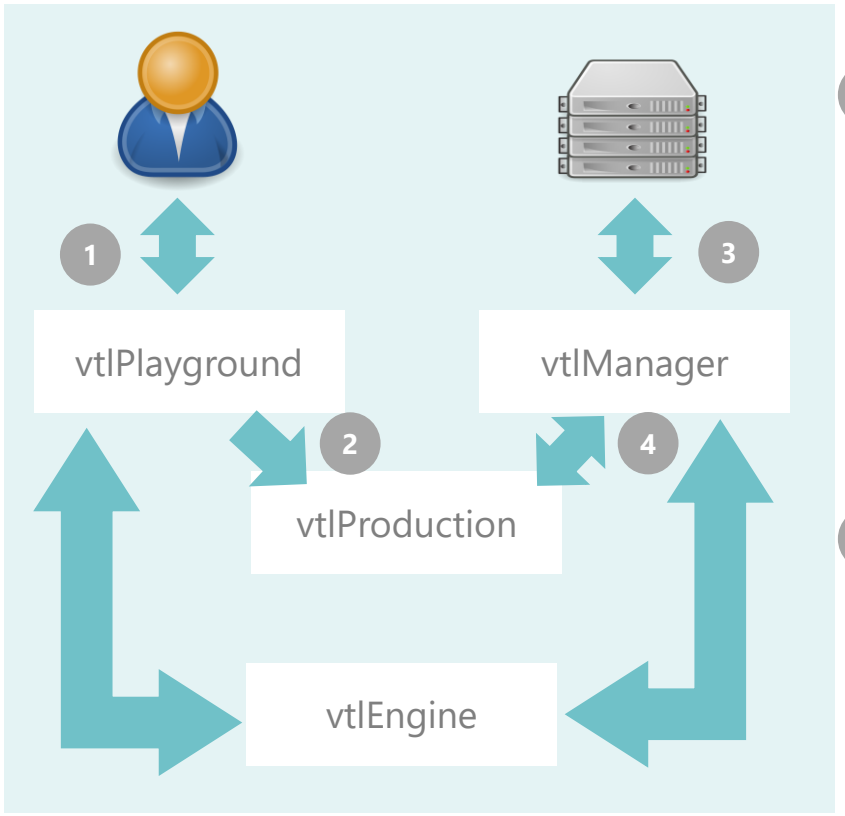


VTL Suite Integration overview

The VTL Suite provides a **web application** for users to write VTL projects, a **REST API to integrate** with other systems, and a task manager to ensure that several concurrent requests can be handled

1. Users collaboratively create VTL scripts, organised in projects, through the vtlPlayground.

2. Coordinators push the project to vtlProduction, which implies that the script is available for execution via the vtlManager API.



3. A Salsa application requests to execute one of the VTL scripts stored in the vtlProduction.

4. The vtlManager takes care of managing the queues and executing the VTL script.

Usable for calculation, transformation

SDMX based

Easy to create own validations

Agenda

Hoe zetten we verdere stappen vooruit in het toepassen van kwaliteitsstandaarden op de productie en publicatie van openbare statistieken?

Wat zijn hier de uitdagingen/valkuilen/opportuniteiten voor de individuele entiteiten en voor het netwerk als geheel?

National Bank OF BELGIUM
Eurosystem

Opportunities

Organizational benefits

Centralization Metadata & Validation Rules

Easy exchange

International Collaboration

Eurostat, BIS, ECB, ILO, IMF,…

World wide community
SDMX/VTL (iso standard)

Process improvements

Structural validation (SDMX)

Reusability

Update / maintain validation

Better data quality

Early data validation (content)

VTL & SDMX compliance

Challenges / Lessons learnt

CREATION

PROJECT TEAM

Project manager, Project coordinator, Business analysts, SDMX expert, Developers, Data process manager, Data manager, Product owner, VTL expert

COLLABORATION

EXTERNAL PARTNERS

Community dependent on new implementations

TIMING

Modelling to SDMX structures

Creation of validation scripts (new language)

NEW SKILLS

Learning curve for SDMX & VTL

New tools to maintain & create SDMX & VTL

Pitfalls

Change in mindset around data organization

Monitor & guard on good practice/structure

Keep quality check on al stages of the process

Thank you

Statistiek Vlaanderen - NBB

21 Maart 2024

frederik.vanhecke@nbb.be

www.nbb.be

National Bank OF BELGIUM

Eurosystem

VTL (Validation & Transformation Language)

- Based on a generic information model that can be used by different standards like SDMX
- Maintained by a VTL Task Force including Eurostat, ECB, ILO, ISTAT , Bank of Italy and INEGI
- Works under the umbrelly of the SDMX Technical Working Group
- Standard language for defining validation and transformation rules
- SDMX compliant
- Allows easy exchange of these VTL rules
- Reusability of rules is key
- Usable by statisticians