

Towards a more efficient data quality process

Collaborating with the sector and with other NCAs

20 October 2021

DeNederlandscheBank

EUROSYSTEEM

Agenda

- Dataloop
- Data Quality Rules

Dataloop – “Digitalize inefficient data quality assessment processes”

Before Dataloop

Data quality analyst



"I feel like I need to click a million times and need five screens to compare two reports"

Supervisory expert



"It is tedious and time-consuming to find out what the status of a report is"

Supervisor



"I don't know where to start my analysis and what has been checked already. And I have no place to post feedback on data quality issues, let alone be informed on the follow-up"

Manager



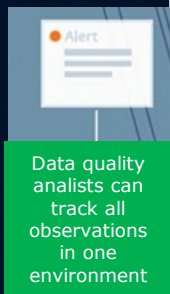
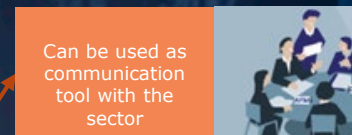
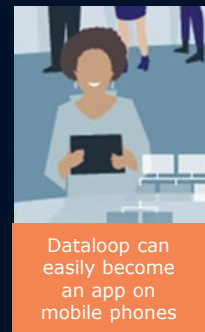
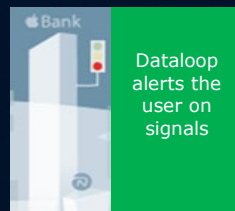
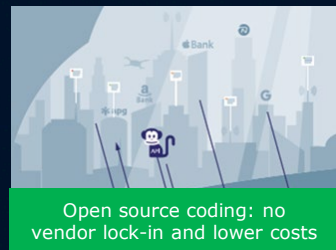
"We need to be able to focus on what we are best in: dealing with the complexity of solving actual data quality issues together with supervisors and the sector"

Benefits of Dataloop for insurance supervision today

- ✓ Centralized access to supervisory data and reports
- ✓ Validation results from different sources are visualized
- ✓ Data quality analysts can track all observations in one environment
- ✓ Live tracking the status of a report

Dataloop as an example of DNB2025 vision in practice

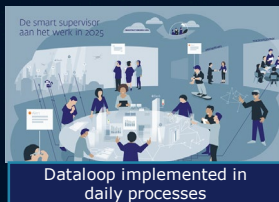
DNBPUBLIC



Technology stack – on the way to implementation

DNBPUBLIC

Azure cloud



DNB Datalake



Milestones so far & what we are going to do

Short-term

iForum pilot: use Dataloop as communication tool with sector

Long-term

Implement machine learning supported signals
Launch Dataloop at another NCA

Proof of Concept

A working proof of concept that is used in the day to day operation of assessing the data quality of **insurance** reports.

POC



2019

Product market fit

Figured out the product market fit resulted in the **data quality journey** with a complete redesign in the user interface. Assessed the fit with the market with an **iForum** initiative, and the fit with the supervisor with the **supervisory signals** pilot.

Product market fit



2020

Scaling ready

Made Dataloop ready for exponential scaling by transitioning from SurfSara to the Azure cloud. Launching customer for both the **DNB Datalake** and the **DNB Kubernetes** cluster.

Scaling ready



2021

Future



>2022

Interested to know more about Dataloop, perhaps even explore the possibilities to implement it in your NCA?

Please send an email to:

Koen van der Veer

Head of Department Supervisory Data on Insurance Companies & Pension Funds

k.j.m.van.der.veer@dnb.nl

&

Ilike Van Beveren

Product Owner of the Insurance reporting team

i.van.beveren@dnb.nl

Agenda

- Dataloop
- Data Quality Rules

Data Quality Rules: project on data quality in collaboration with sector

Goal: data-driven supervision requires high data quality



More efficiency

Efficient internal data quality process at entities.



Better rules

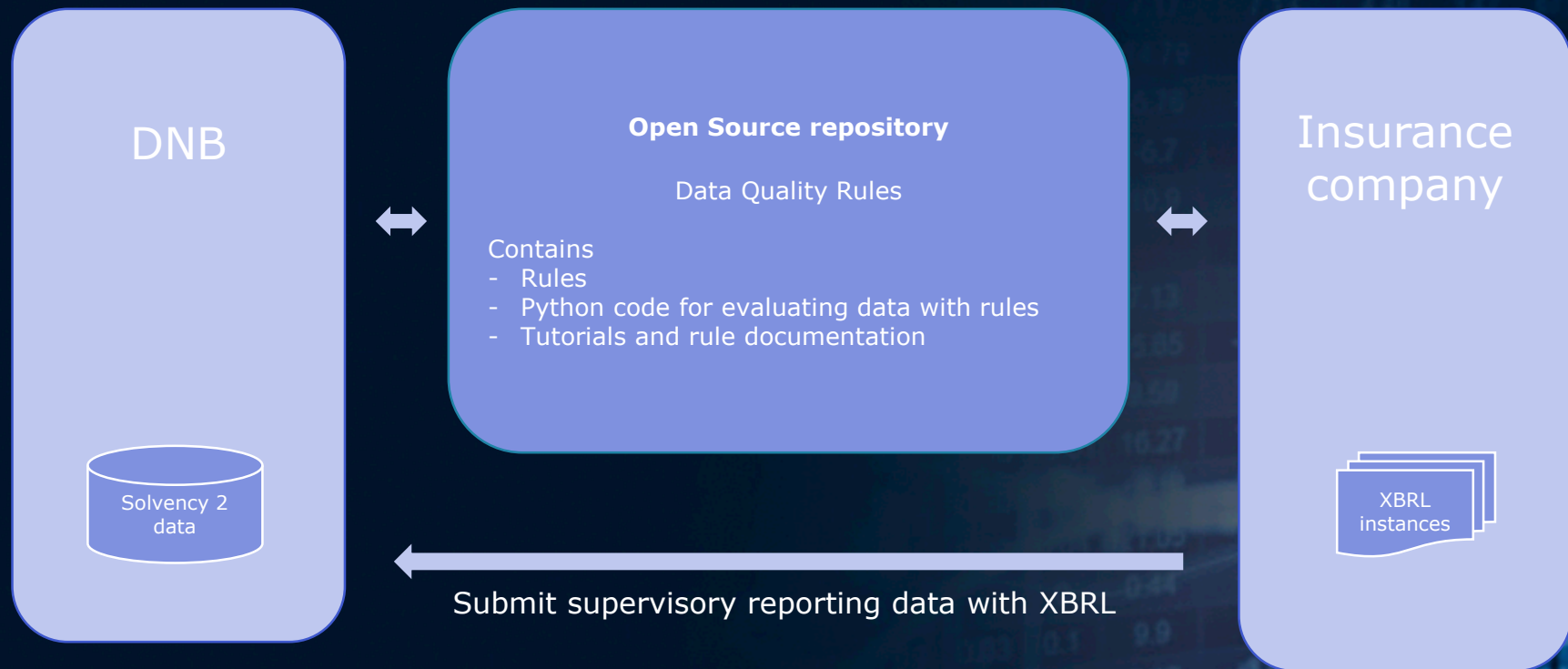
Extensive set of validation rules that go beyond what is possible in XBRL.



Collaboration

Broad access to digital environment with validation rules for DNB and entities, with continuous development.

Design Pilot: digital platform




Design Pilot: digital platform (2)

- We use **Github** as digital platform
Platform for sharing code. Contains strict access control, collaboration functions, forums, to-do lists
- Publication **Data Quality Rules** repository
github.com/DeNederlandscheBank
- Open Source Software

The Open Source Software of DNB is available under license MIT/X of EUPL

Design Pilot: digital platform (3)





De Nederlandsche Bank
EUROSISTEEM


De Nederlandsche Bank


Open Source projects of DNB

Amsterdam <https://www.dnb.nl/>


 **Repositories** 5

 **Packages**


 **People** 2


 **Projects**

Pinned repositories



 **solvency2-data-science**


Data science projects with Solvency 2 data

 Jupyter Notebook



 **data-patterns**


Package for generating and evaluating patterns in quantitative reports

 Python  3



 **dnb-nlp**


Experimental natural language processing projects with central bank and supervisory documents

 Python  1

 **solvency2-rules**

Solvency 2 Data Quality Rules and Patterns







 Jupyter Notebook  1

 wjwillems name FTK tables changed c54989a 2 days ago 165 commits

eggs	new unittest	16 days ago
data	adjustment FTK notebook	2 days ago
docs	taxonomy added	7 months ago
ftk-rules	adjustment FTK notebook	2 days ago
notebooks	final adjustments notebooks	2 days ago
pkgs	pkgs updated	23 days ago
results	update	2 months ago
solvency2-rules	new rules + demo files + tutorial	5 days ago
src	Merge branch 'master' of https://github.com/DeNederlandscheBank/dat...	2 days ago
tests	adjustments demo file	2 days ago
.gitattributes	initial commit	8 months ago
.gitignore	tox and build deleted from repo	21 days ago
.travis.yml	tox changes	21 days ago
HISTORY.rst	name FTK tables changed	2 days ago
LICENSE	license updated	2 months ago
MANIFEST.in	tox changes	21 days ago
README.rst	datapoint defs added	9 days ago
setup.py	v0.2 -> v0.3	2 days ago
tox.ini	tox changes	21 days ago

README.rst

data-quality-rules

 release  v0.3.0  build  passing  license  MIT/X

This is Data Quality Rules repository for the Solvency 2 and the Financial Assessment Framework (FTK) quantitative reporting templates (QRTs). In this repository De Nederlandsche Bank publishes additional validation rules to improve the data quality of the QRTs.

This repository is part of the iForum pilot 'Data Quality Rules'. Want to know more? Please contact P.M.Willems@dnb.nl

- Free software: MIT/X license

Features

Here is what the package contains:

Design pilot: types validation rules

DNB PUBLIC

- Validation rules in the **taxonomy** (maintained by EIOPA)
- Additional Solvency 2 validation rules by DNB-Statistiek in Excel
- New Solvency 2 validation rules
 - granular data derivatives
 - granular data assets
 - comparison data of two periods

Design pilot: automatic generation of new validation rules

DNB PUBLIC

- To check the data quality of the data, we need a lot of validation rules. We generate these using an expression, which the algorithm uses to find the rules.
- These rules check whether the ISIN code (unique asset ID code) has the right NACE (economic sector issuer) connected to it. A lot of the ISIN codes had the wrong NACE connected to it.
- The rules are checked by experts before they are published.



Interested to know more about Data Quality Rules?

Please send an email to:

Annick van Ool

Supervision specialist

Insurance supervision - Data-analytics expertise centre

a.w.m.van.ool@dnb.nl