



KNIME DataHop

How ING Uses KNIME to Reduce Complexity in its Audit Process

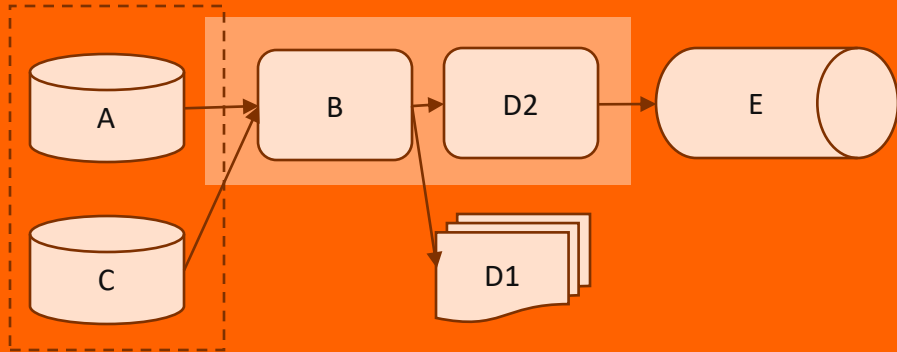
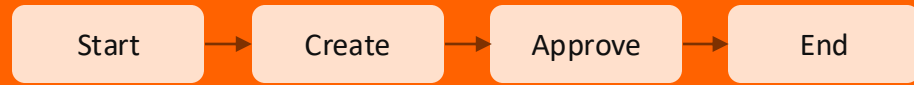
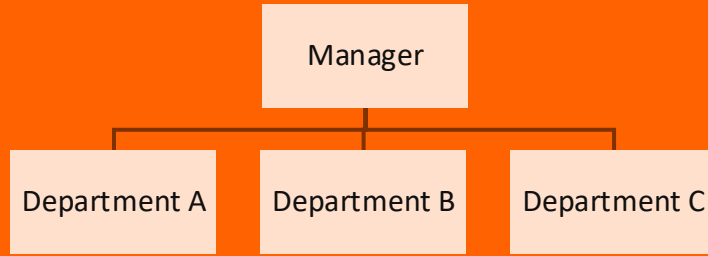
July 1, 2025

Matheus Toscano / Tjasse Biewenga

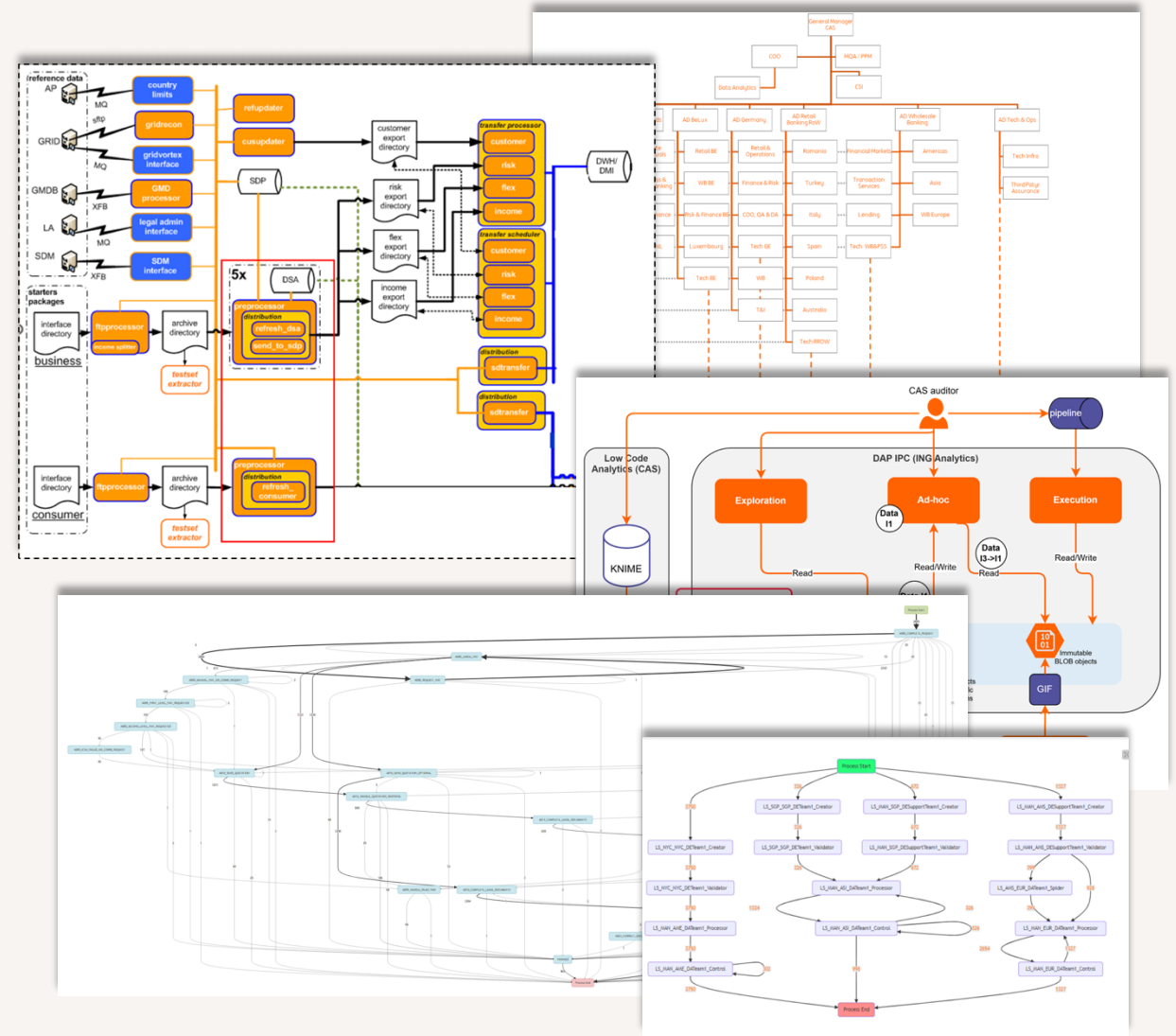


do your thing

Picture



Reality



ING

ING is a global financial institution with a strong European base, offering **retail and wholesale banking** services to customers in more than 100 countries.

+ 60.000
employees

+ 40 million
customers

10.079 organisation entities

1.445 job profiles

3.575 business IT applications

97.138 controls



Internal Audit

+ 400 Staff

44 teams

20 countries

Audit Universe

1.288 auditable entities

400 audits per year

6 – 10 weeks lead time

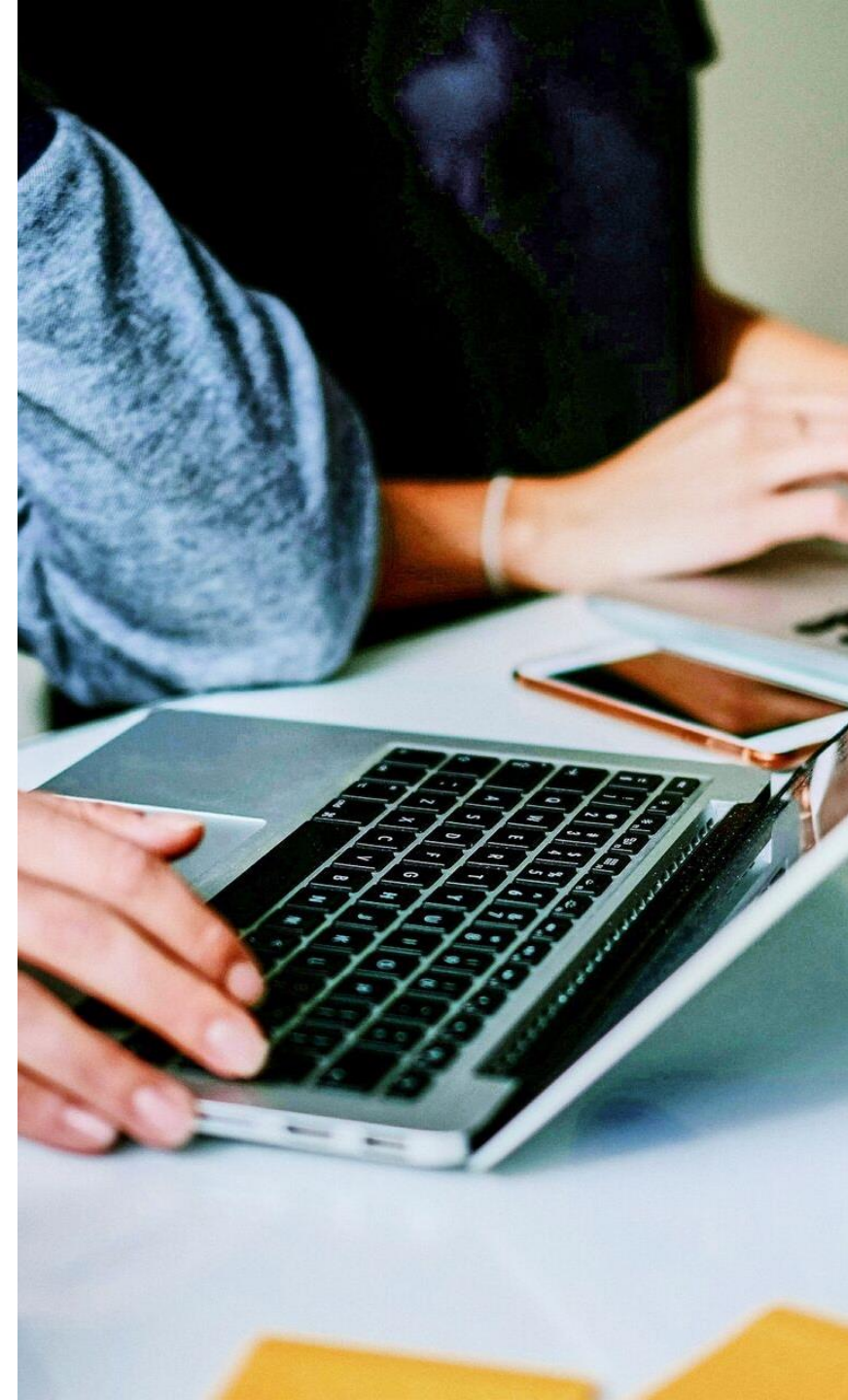
Our role

“ to provide an independent assessment of the design and effectiveness of internal controls”

Analytics Strategy

Analytics for all

Flexibility, re-useability



The Audit Process

Audit Planning



Preparation



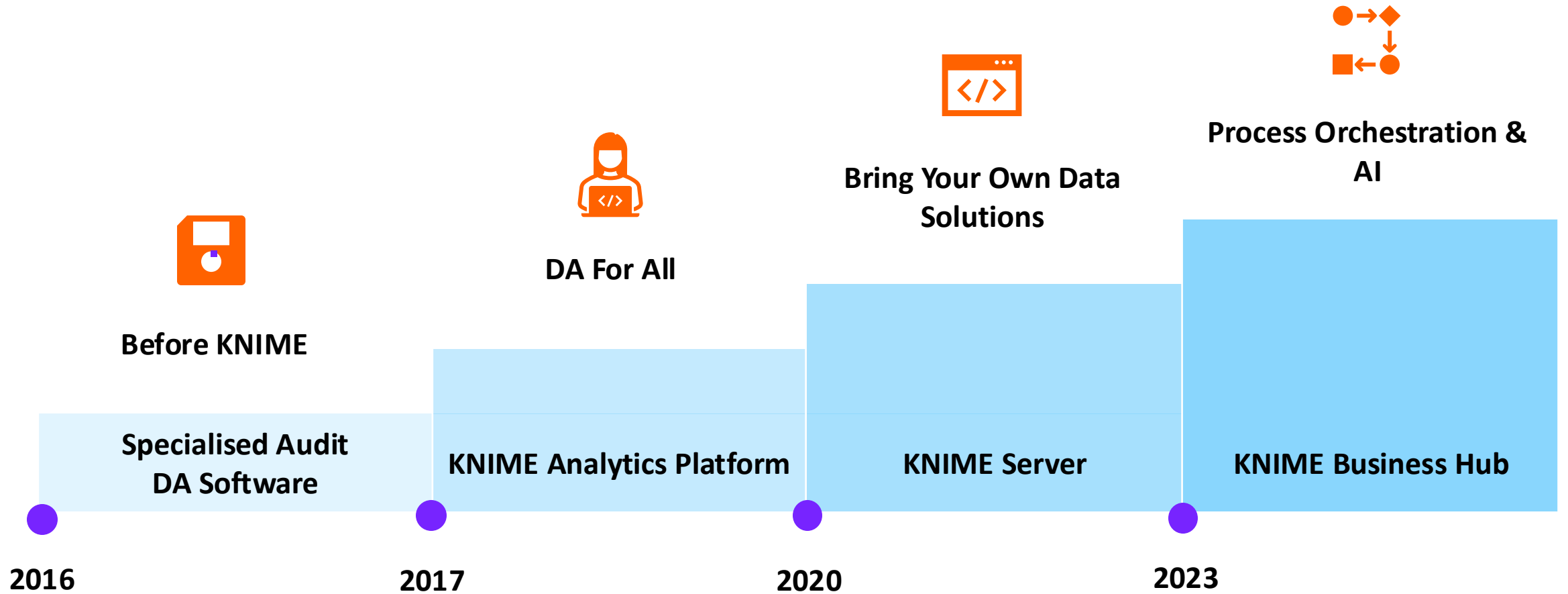
Testing



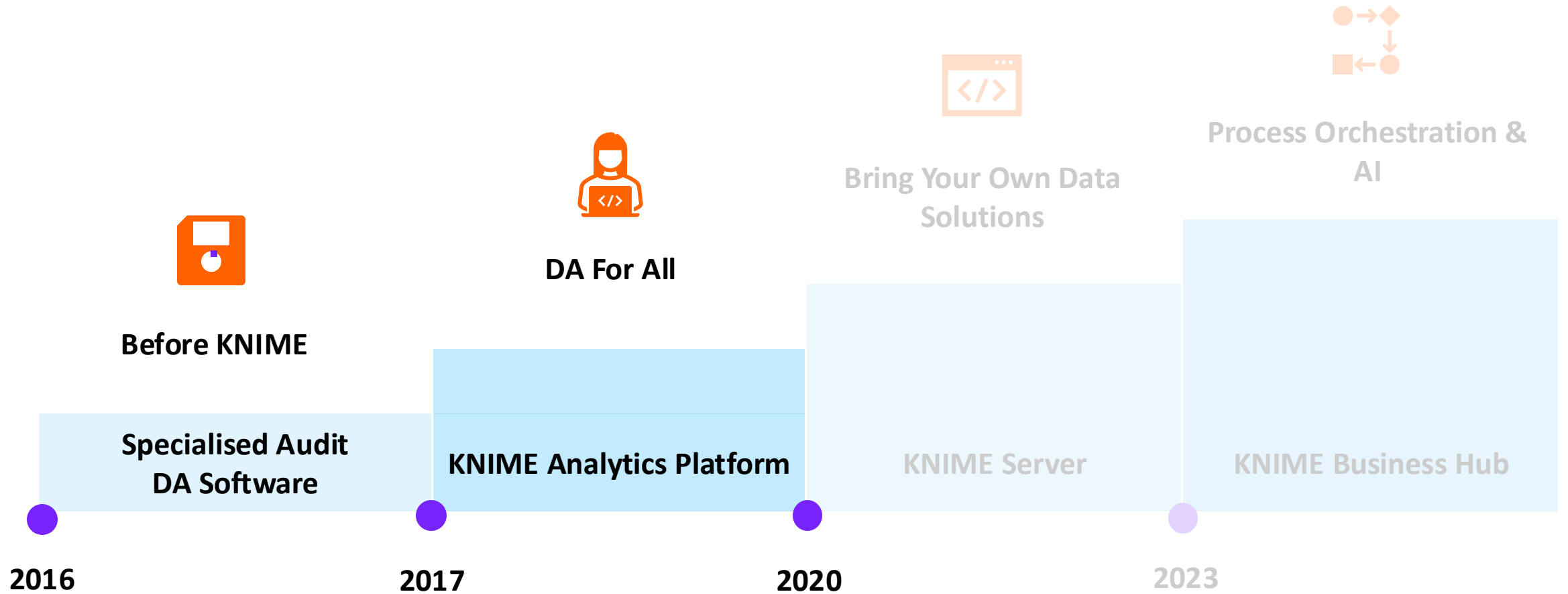
Reporting



The KNIME journey

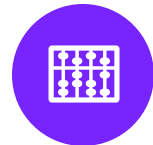


Streamlining analytics tooling



Streamlining analytics tooling

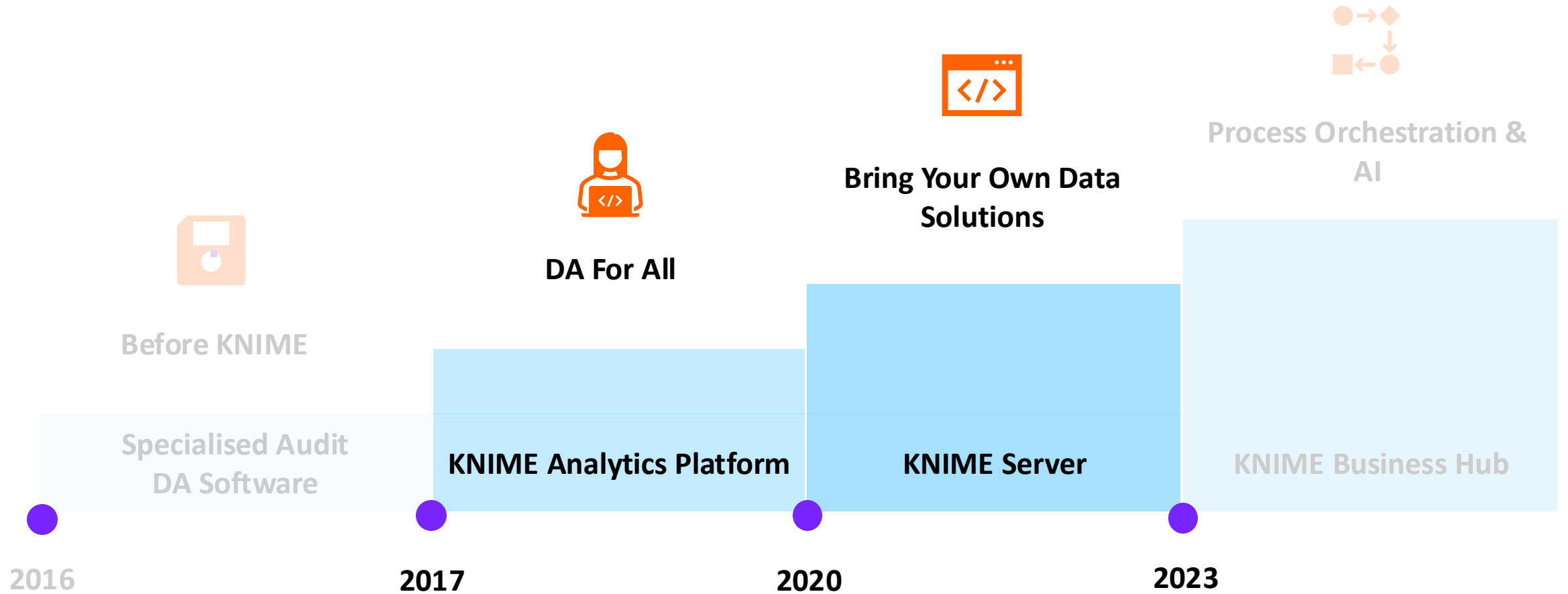
Doing more with less



Before KNIME Analytics Platform	After KNIME Analytics Platform
Different tools in different teams and countries	One standard tool for the department (other tools allowed)
Manual tasks, very limited reusability of logic	Reusability by default, consistent audit approach, collaboration
Restricted capabilities and data volumes	Different data types, database connections
License costs	No license costs

Self-service Audit Analytics

Essential audit analyses as web pages



Self-service Audit Analytics

Essential audit analyses as web pages



Outlier Detection

Outlier Detection

The current solution utilizes a set of different algorithms to identify potential outliers in a dataset.

An overview of the number of outliers detected per algorithm are presented on the top.

The ultimate outlier count shows the number of outliers that are detected by all used algorithms (if DBSCAN is disabled both DBSCAN outliers and ultimate outliers will be set to 0).

The graphs and tables below present in more detail the outliers per algorithm.

Ultimate Outliers

0

Mean Deviation

0

DB SCAN

0

Isolation Forest

1



Process Mining

Workflow

Process Mining 2.0

Generic Solutions

Version: First Hub Versi... - Latest, created on Feb 19, 2025 9:40 AM

Versions

Data Upload
Widget for user to upload a system log file.

Column Selection Screen
Prompts user to inform the columns for

- process id (mandatory)
- event (or status, mandatory)
- timestamp (mandatory)
- user (optional)
- process group (optional)

Column Renaming
Rename the data columns according to the values informing so all the data manipulation can be made on standard names.

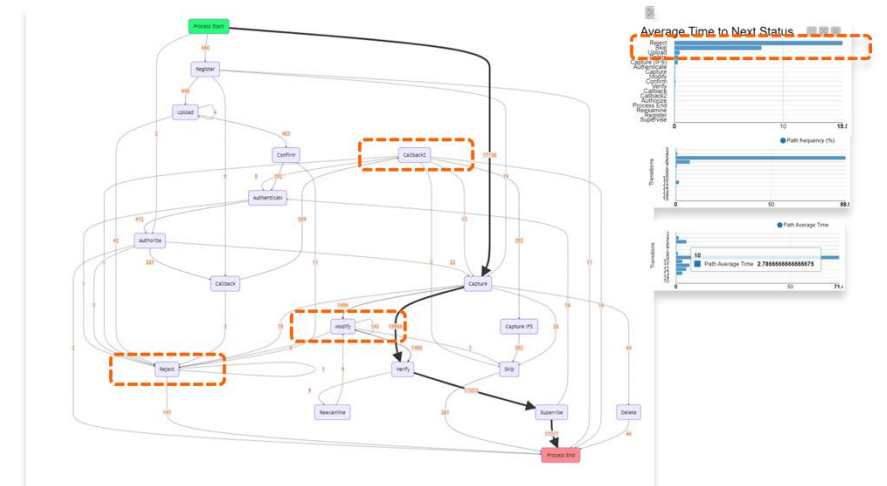
Pre-processing
Generates in advance information used for the process visualisation in insights. Two datasets:

- Overview of all process ids and respective indicators
- System log with additional column fields

Process Mining View
Include the visualisations for menu / parameter filtering and selected insights.

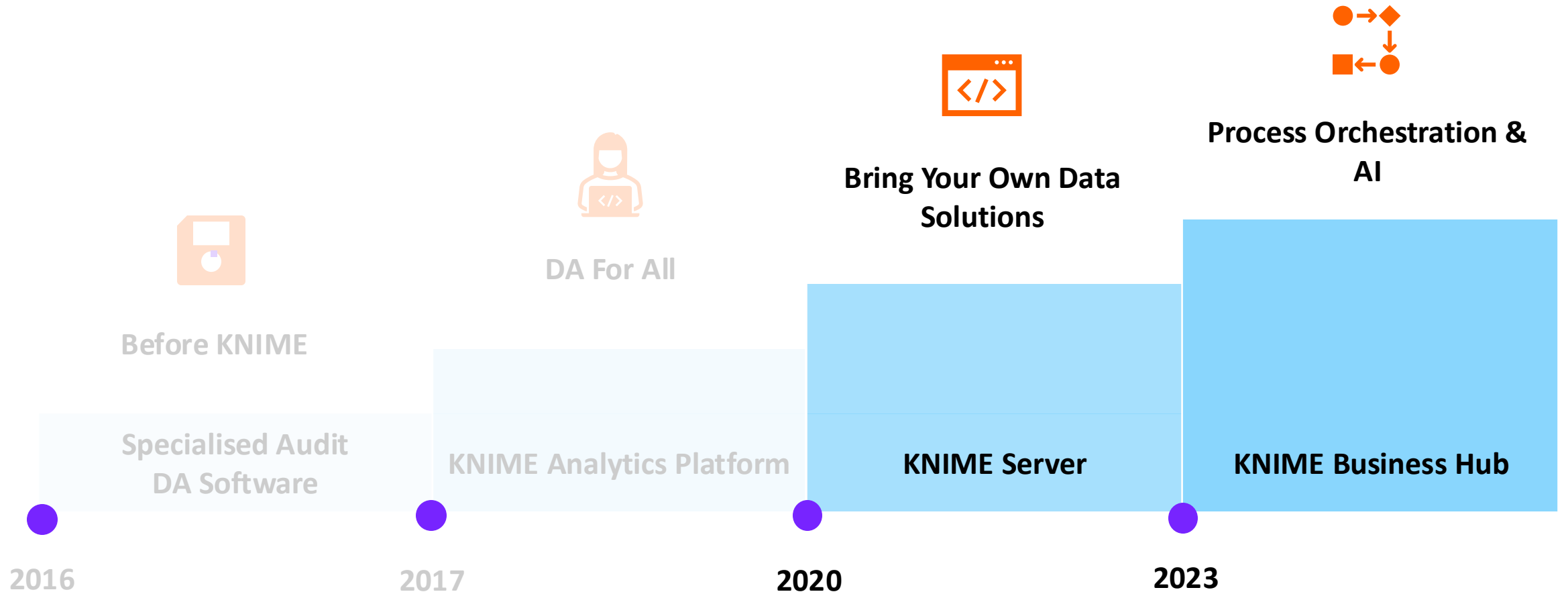
File Upload Column Selection Process Column Renaming Pre-processing Process Mining

output: system log output: renamed system log top: process overview bottom: system log



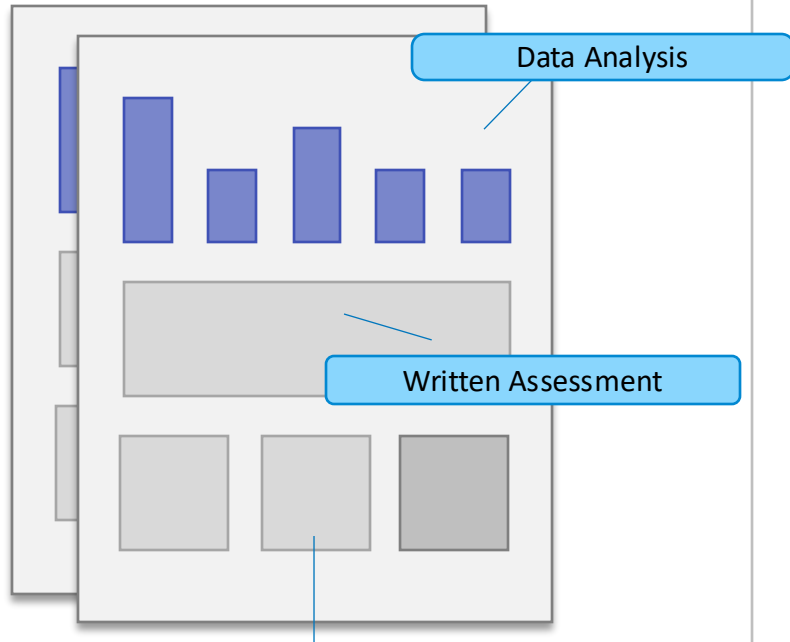
Audit Process Automation

Reducing manual work in preparing documents




Audit Process Automation

Reducing manual work in preparing documents



Conclusion (scope impact)




Pre-populated template?

Phase 1

Empty template based on:

- Audit Information
- Sections to be filled

Audit Info Reader



Field	Value
String	String
Auditable Entity	AE XXXX
Audit Name	Audit of XXXXX
Team	Retail NL
Category	Regulatory

Phase 2

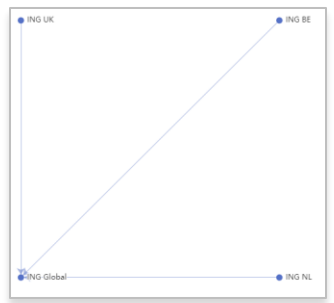
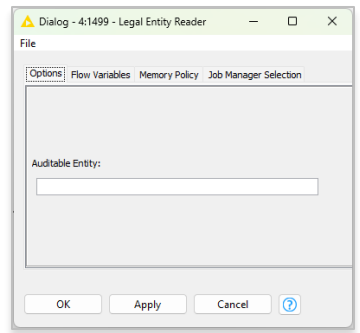
Automate individual sections:

- Source the data used as input
- Prepare charts and tables

Legal Entity Reader



Legal Entity Viewer



Phase 3

Use of Gen AI to:

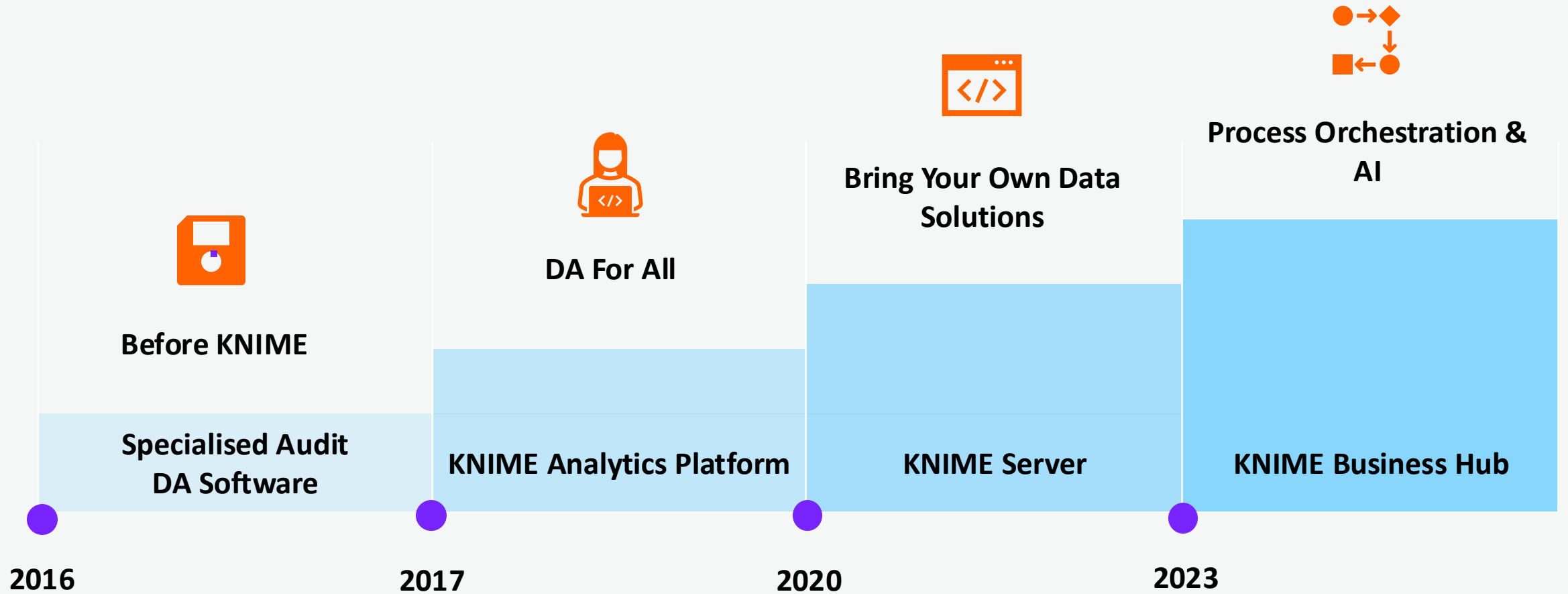
- Include assessments based on previous work
- Common reporting language



Python Script



The KNIME journey



Lessons Learned in our journey

Use it or Lose it



- Training only is not sufficient
- Create opportunities for usage (e.g. hackathons)
- New ideas from power users

Divide & Conquer



- Divide complex solutions into modules
- Create reusable components
- Mature the process through continuous improvement

Look for Value Added



- Balance what is possible with what is needed
- Symptoms:
 - Repetitive tasks
 - Limited IT systems



do your thing

Thank you!

www.ing.com