

# Process mining and data analytics in P2P data

Ove Haugland Jakobsen

Senior Data Scientist, OAGN Innovation Lab

# Benefits of data analytics

- Centralised team to conduct data processing and analytics
- Controls performed on 100 percent of POs – no audit sample
- Fully automated controls – far less time-consuming
- Resuable code – the same code can be used across audit assignments and fiscal years
- Fully reproducible results – everything is documented in code (using SQL Notebooks and Rmarkdown/Quarto)



# P2P data and problem definition

## **Purpose**

- Check four-eyes principle (separation of duties) for *all* POs
- Check if all POs conform to our expectations for the process

## **Data**

- Data from The Norwegian Agency for Public and Financial Management covering more than 60 agencies
- Workflow (log) data for more than 630k purchase orders (POs)

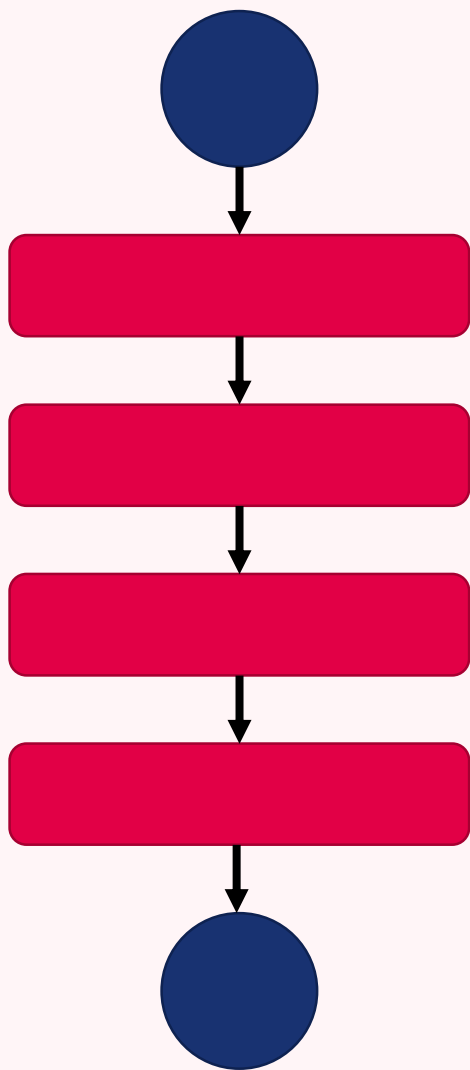
## **Problem**

- Transforming log data to event data and identify relevant events

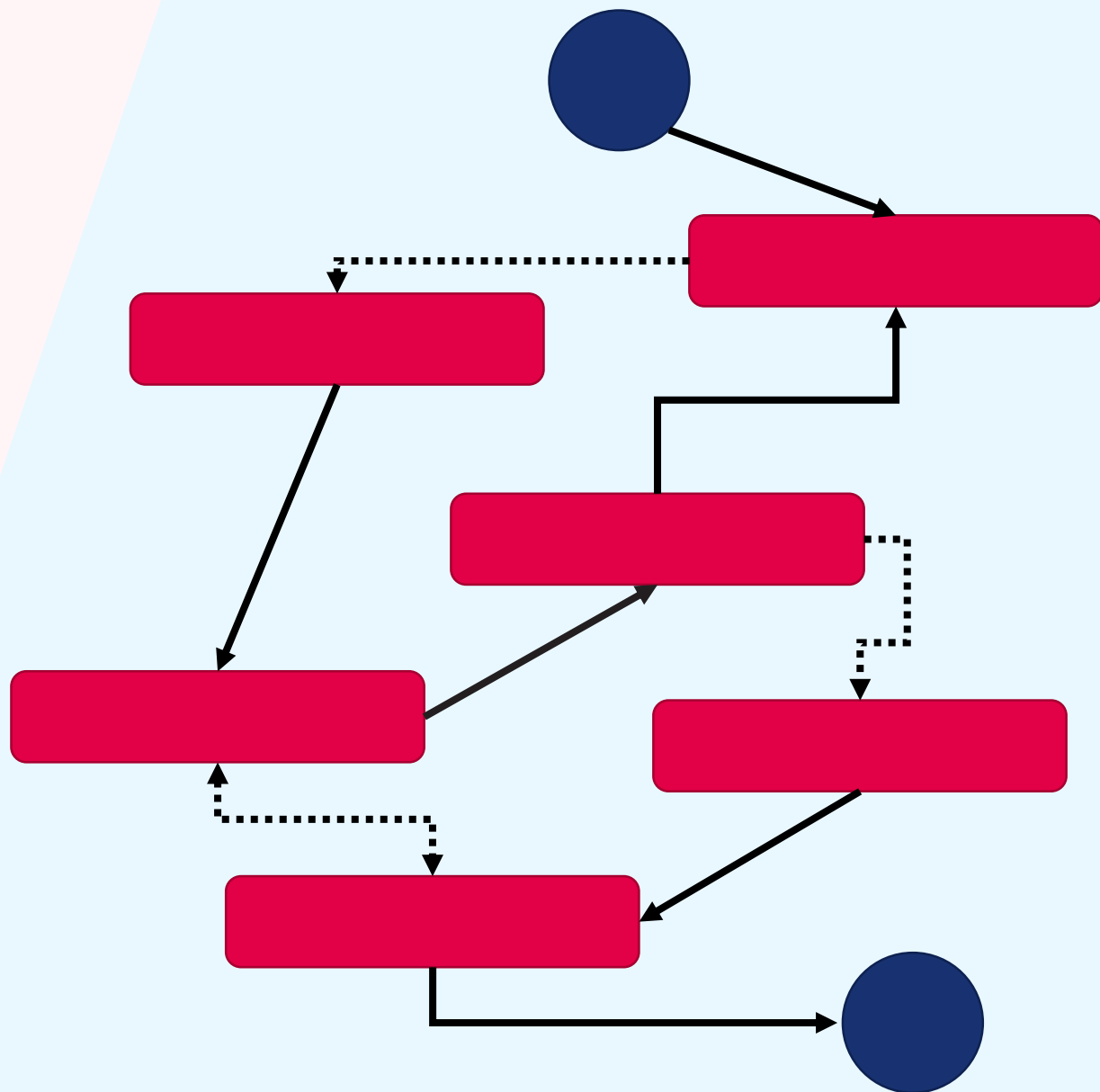
# Using process mining in data analytics

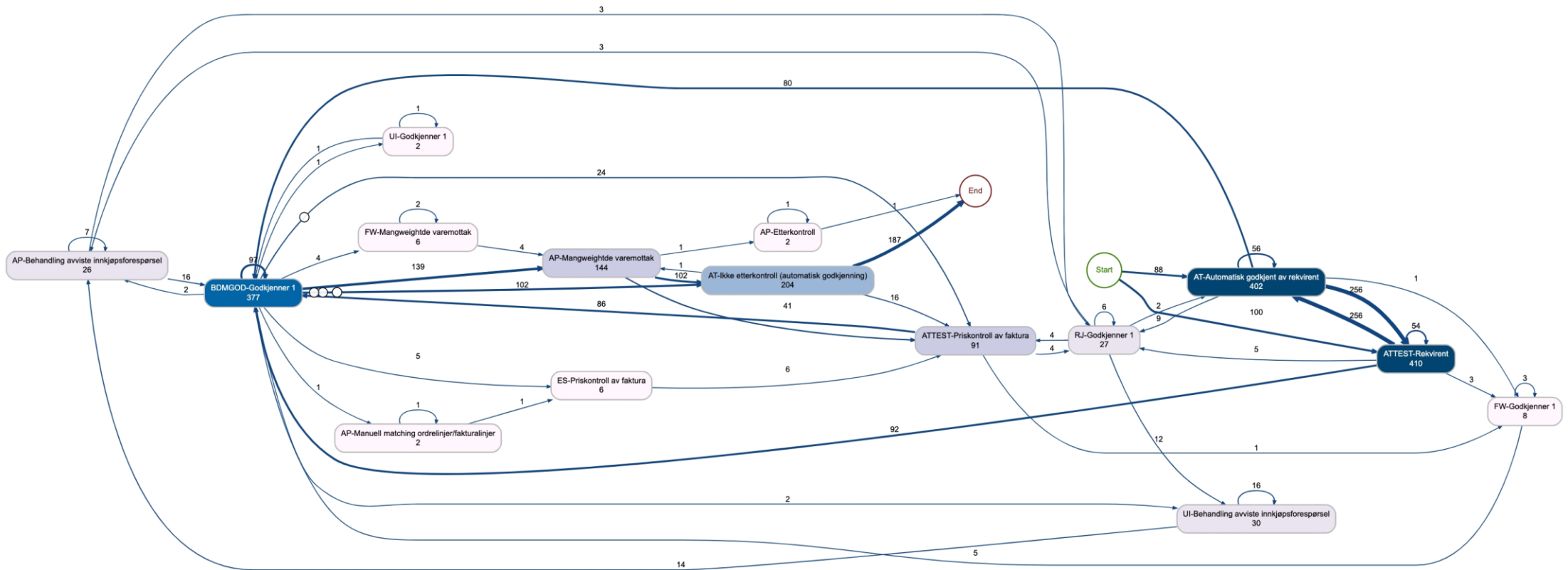
- Process mining is a tool for processing analysing data related to processes
- Recreate the actual process flow through *event data*
- An analysis of what *actually* occurs in a process

Theory



Data





# Results

- 630k purchase orders for FY2021 with > 53 BNOK in total value
- 148 POs with breach of separation of duties, totalling 2 MNOK
- For FY2022 we expect a 5-fold increase in POs (more agencies added)
- More controls to be added for FY2022 based on experiences from FY2021

# Success factors

- Centralised agency for government accounts in Norway
- Similar data structures across agencies – most agencies are moving to the same platform – code can be reused for each agency
- Cross-disciplinary collaboration within OAGN: data science, IT, financial audit resources pooled together



# Tools

- Microsoft SQL server
- R and RStudio
- tidyverse for data wrangling/transformation using R
- bupaR for process mining using R
- Rmarkdown (Quarto from FY2022) for documentation

