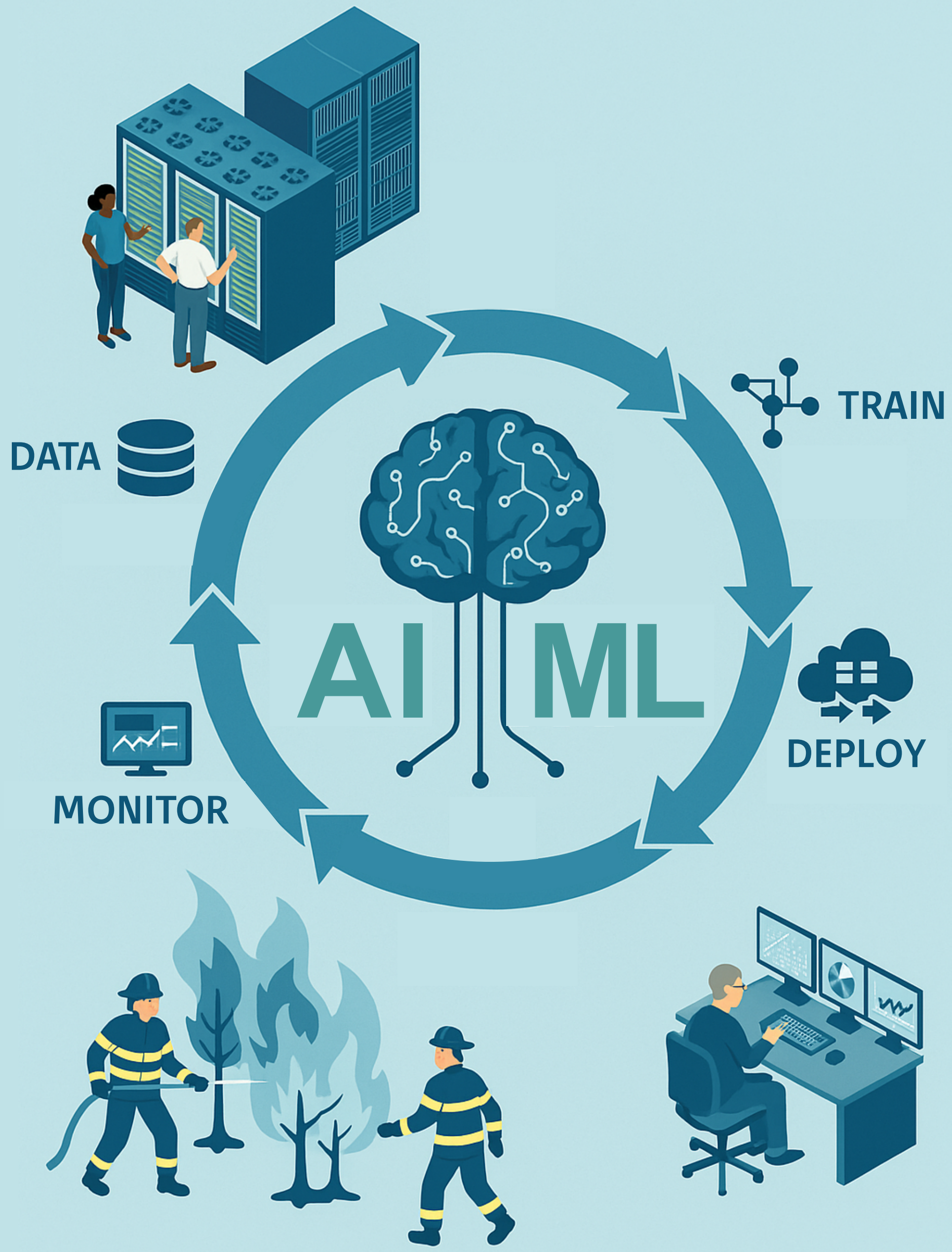


Royal Netherlands
Meteorological Institute
*Ministry of Infrastructure
and Water Management*

MLOps practices at KNMI

The Collaborative Quantitative Impact Forecasting use case

EMS2025

Alice Alfonsi
Jasper van Nieuwenhuizen
Rosina Derks
Margo van der Stam
Luca Trani
Jouke de Baar

Transitioning AI/ML research to operation is crucial to deliver reliable services and decision support that safeguard society

Research

- Innovation, experimentation
- Data (somewhat) available
- Proof of concept
  Model performance
  Novelty
- Publications, Reports

Operation

- Robustness, reliability
- Established data streams
- Explainability and transparency
  Flexible model serving
  Continuous learning
- Operational product

KNMI MLOps initiative

- Official dedicated agile team since 2025
  4 developers, 1 scrum master, 1 product owner
- Goals
  - Operationalise AI/ML products
  - Develop and maintain AI/ML platform for researchers
  - Provide guidelines and promote best practices
  - Foster collaboration and exchange knowledge

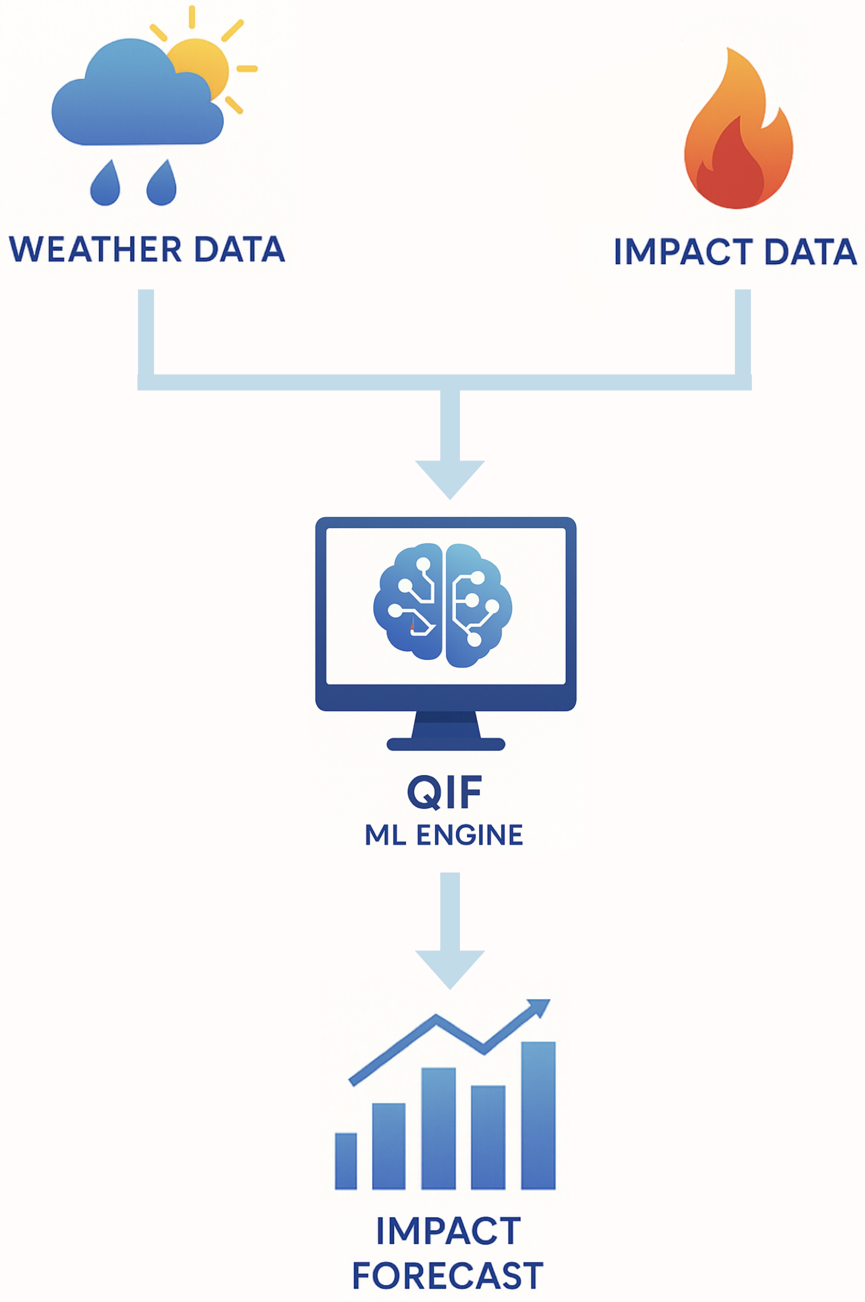
> “
>
> AI/ML to better understand impacts of extreme and compound events thus improving our communication and information capabilities

KNMI
AMBITION 2035

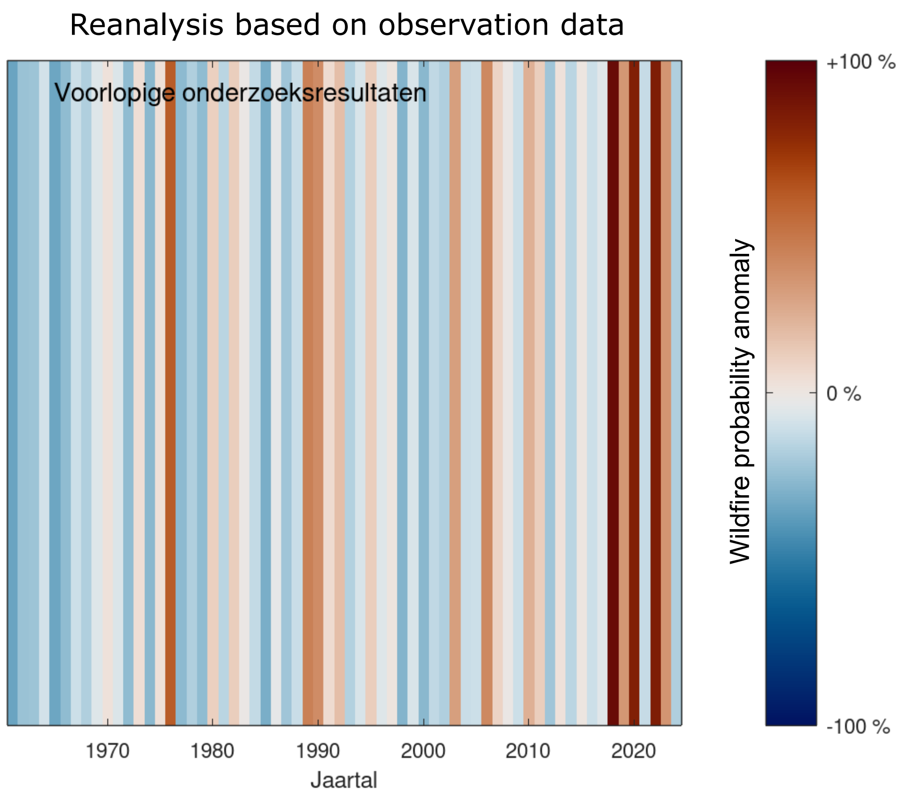
C-QIF for wildfires

Cooperative quantitative impact forecasting

- Data-driven framework relying on historical weather and wildfire records
- Delivers probabilistic daily forecast of wildfire occurrence in the Netherlands
- Cloud-based, real-time operations for scalable, continuous data processing
- Supports operational resource allocation and public communication
- Stakeholders: KNMI weather room, NIPV



Wildfires in The Netherlands



- Wildfire occurrence and intensity are rising
- Assessing impact is complex due to many factors at play
  - High number of wildfires anticipates severe impact
  - Time is crucial in preventing high impact
- Urgency to shift from reactive to proactive approach

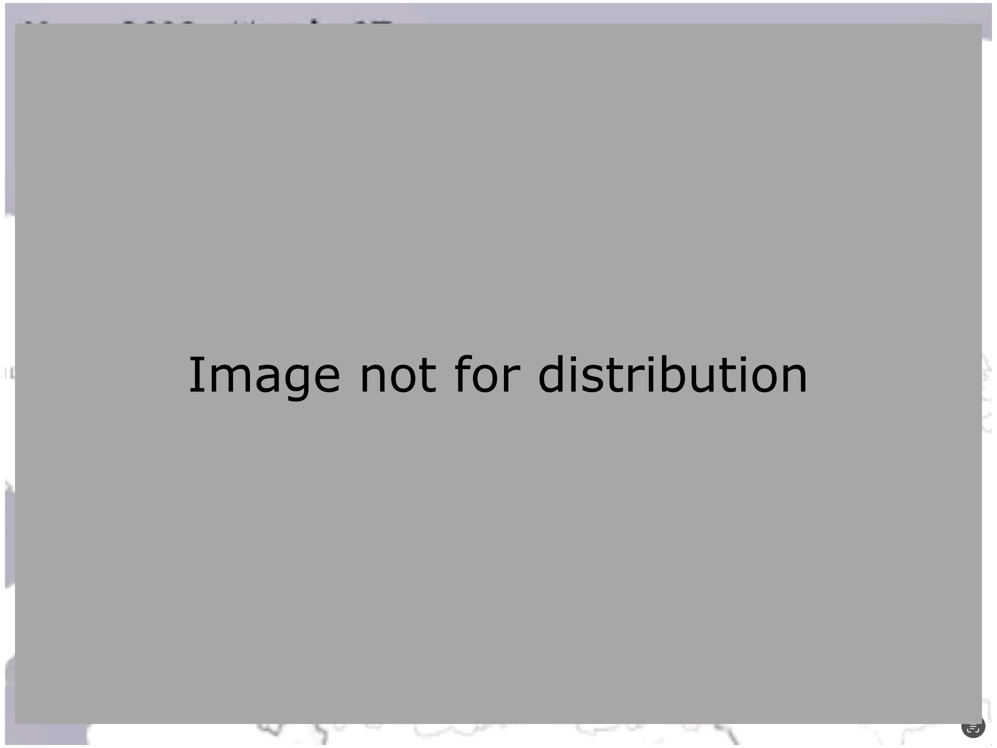
MVP: observations-based

Based on KNMI De Bilt observations and NIPV number of wildfires dataset

- All fires classified as wildfire from 2017
- Daily mean windspeed and relative humidity, min/max temperature, precipitation, month and averages over 14-days

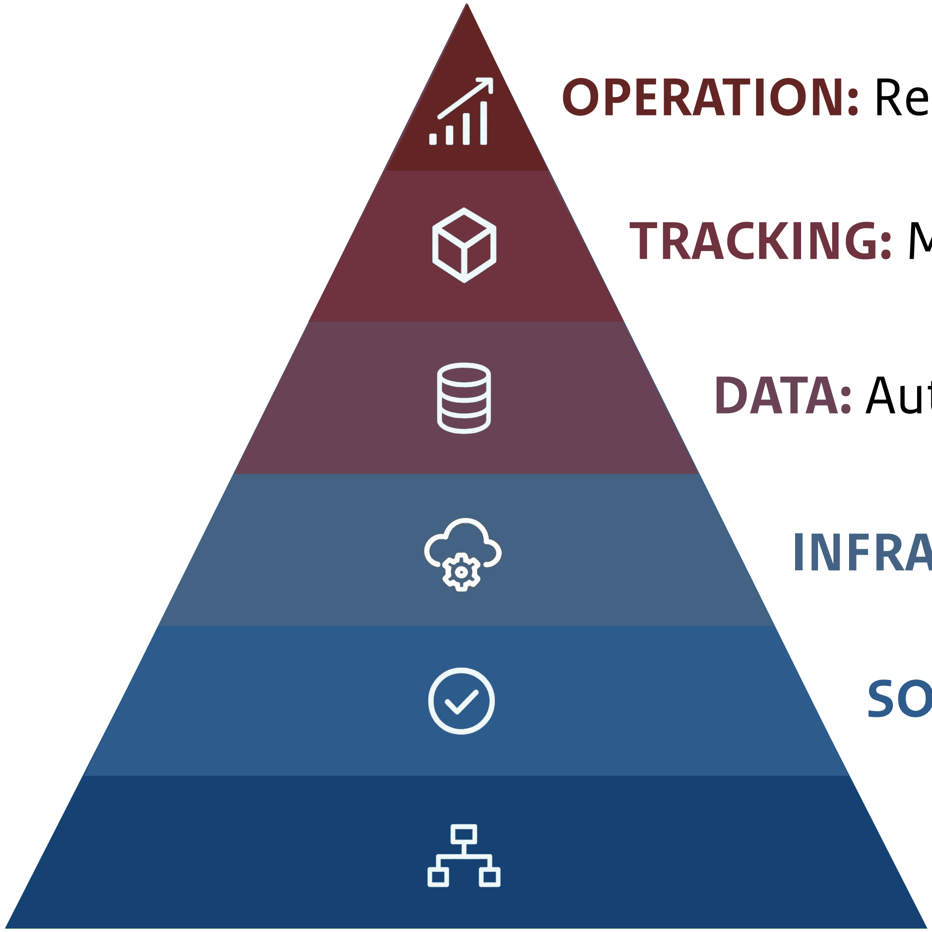
Focus

- Test the full base infrastructure early
- Provide hindcast validation of the model



Around 750-1000 incidents per year

Approach building blocks



OPERATION: Real-time dashboard, monitoring, secure and maintainable model

TRACKING: Model development and data tracking in MLFlow, reproducibility

DATA: Automated data stream, preprocessing and storage

INFRASTRUCTURE: Scalable and secure AWS cloud, CI/CD pipelines

SOFTWARE PRACTICES: Clean code, documentation, testing

GOVERNANCE: Guidelines, requirements, continuous feedback

Technical timeline

MARCH 2025
RESEARCH PROTOTYPE

- Local Octave code
- Historical wildfire records + KNMI observations/ forecast
- M5' ensemble model

- Port to Python & Evaluate model alternatives
- Identify required data and sources, processing and storage
- Build infrastructure and pipelines
- Write documentation

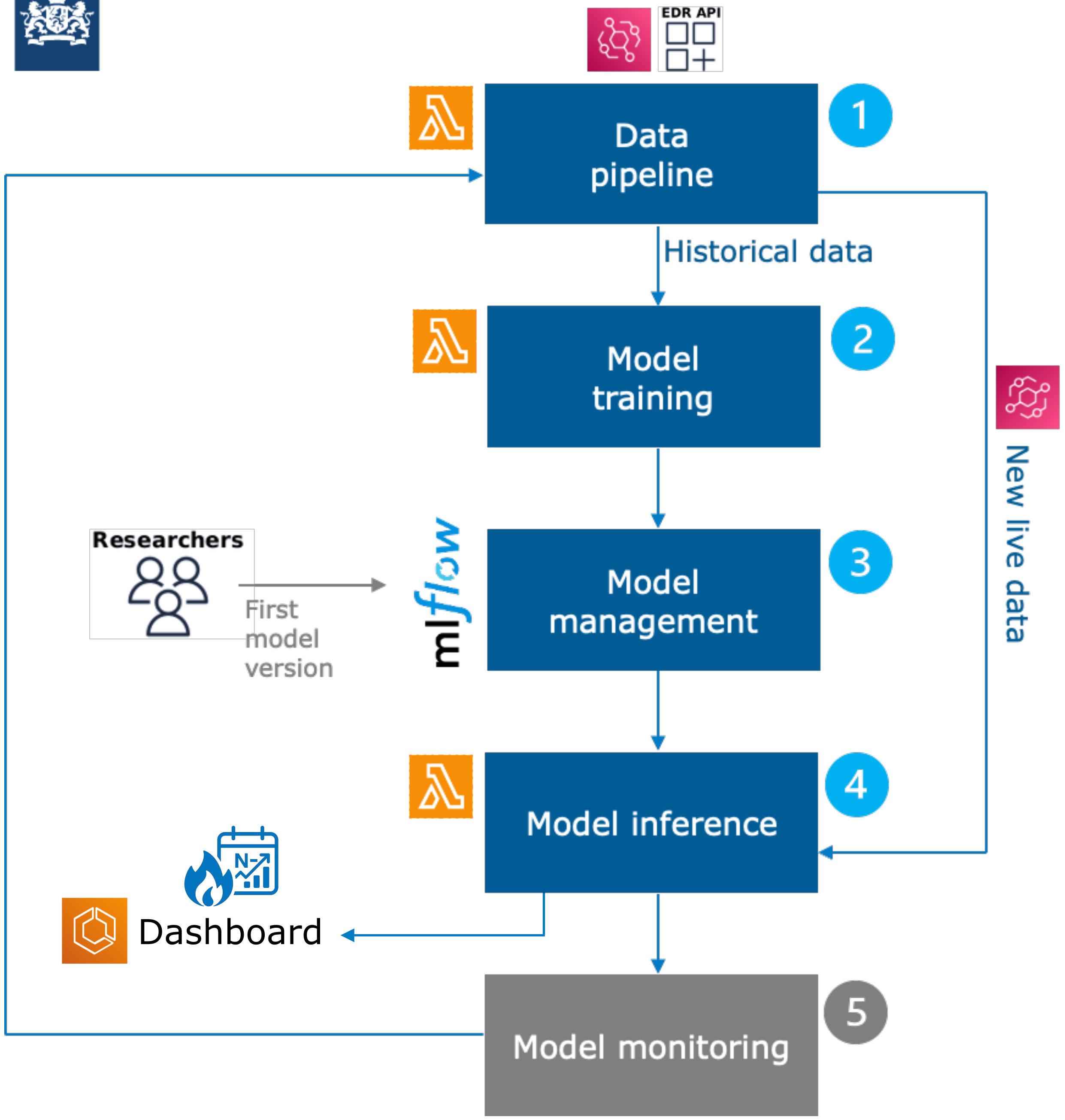
AUGUST 2025
OBSERVATIONS MVP

- Cloud-based python
- Historical wildfire records + KNMI observations
- XGBoost QuantileRegressor model

OPERATIONAL FORECAST

Operational workflow

- Automated model tracking in self-hosted MLflow, manual version promotion
- Automated historical and live observation data retrieval and processing
- Input and output data stored as parquet dataset partitioned by date
- Manual model monitoring and training
- Automated daily inference results with incoming live data as dashboard service



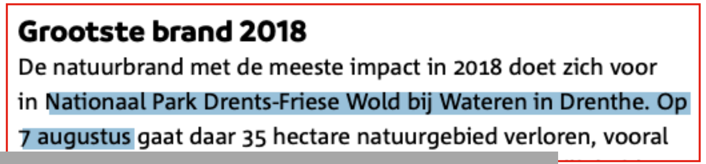
Hindcast dashboard

SCREEN CAPTURE
NOT ALLOWED

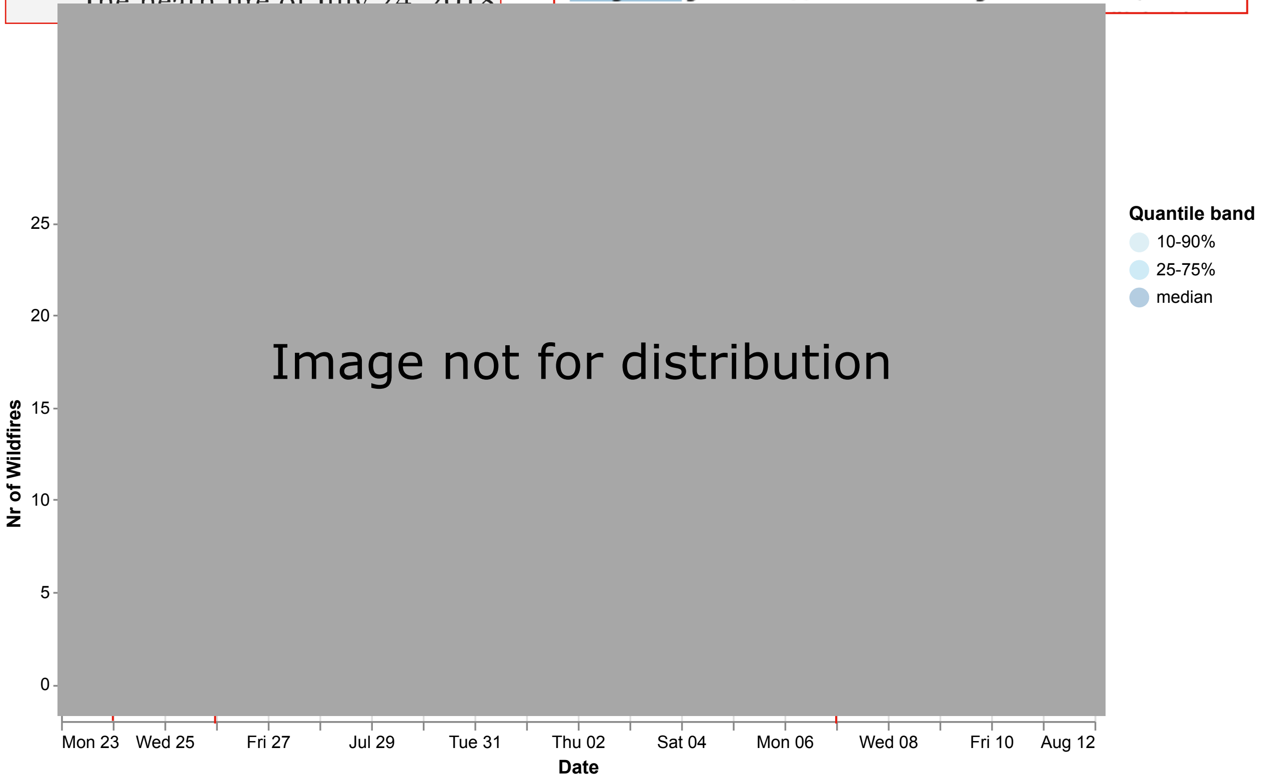

stichting het nationale park
de hoge veluwe

ABOUT THE PARK / FOR THE PRESS / PRESS RELEASES

The heath fire of July 24, 2018

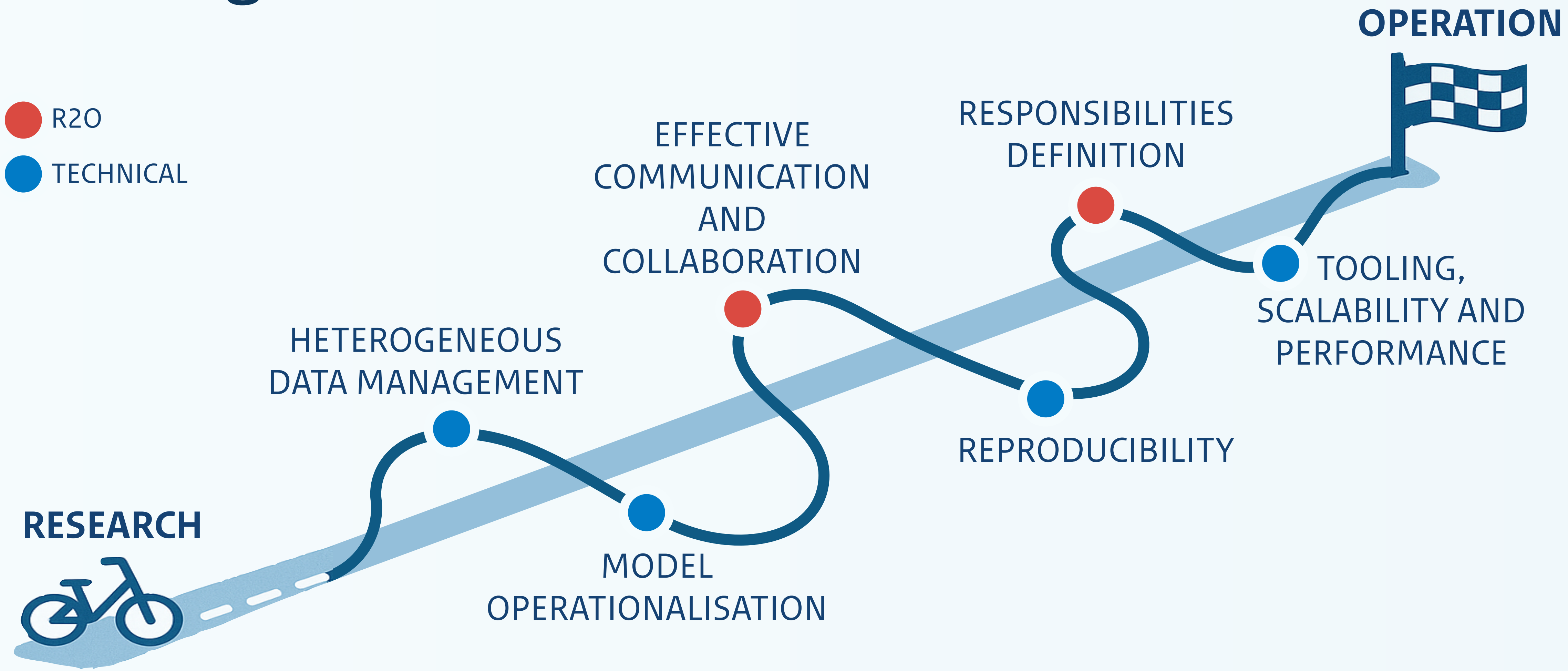

Grootste brand 2018

De natuurbrand met de meeste impact in 2018 doet zich voor in Nationaal Park Drents-Friese Wold bij Wateren in Drenthe. Op 7 augustus gaat daar 35 hectare natuurgebied verloren, vooral

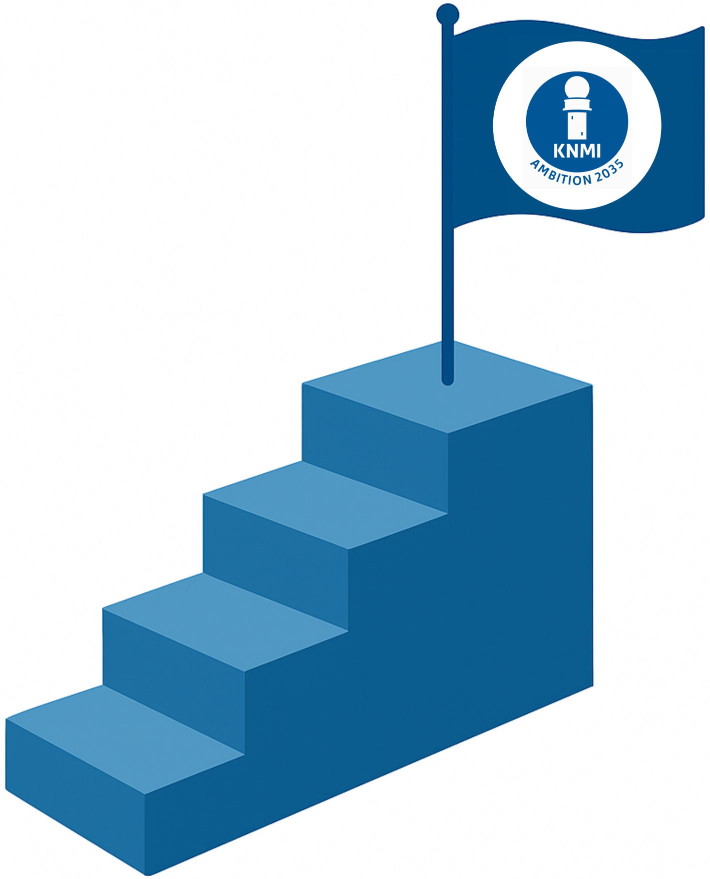


- Operational framework with observation data in place
- First hindcast results reflect expectations
- Model research and validation ongoing

Challenges and lessons



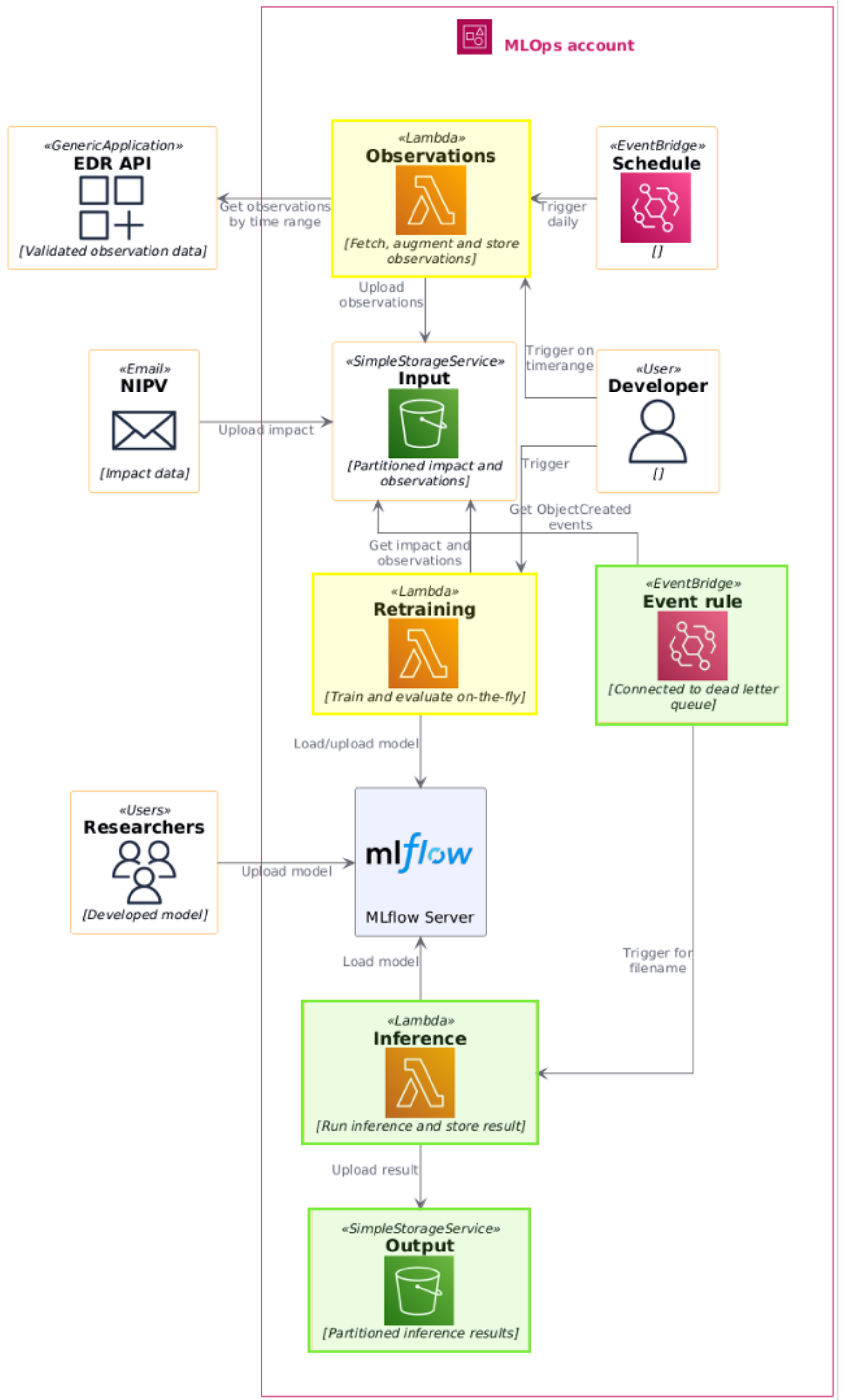
Outlook



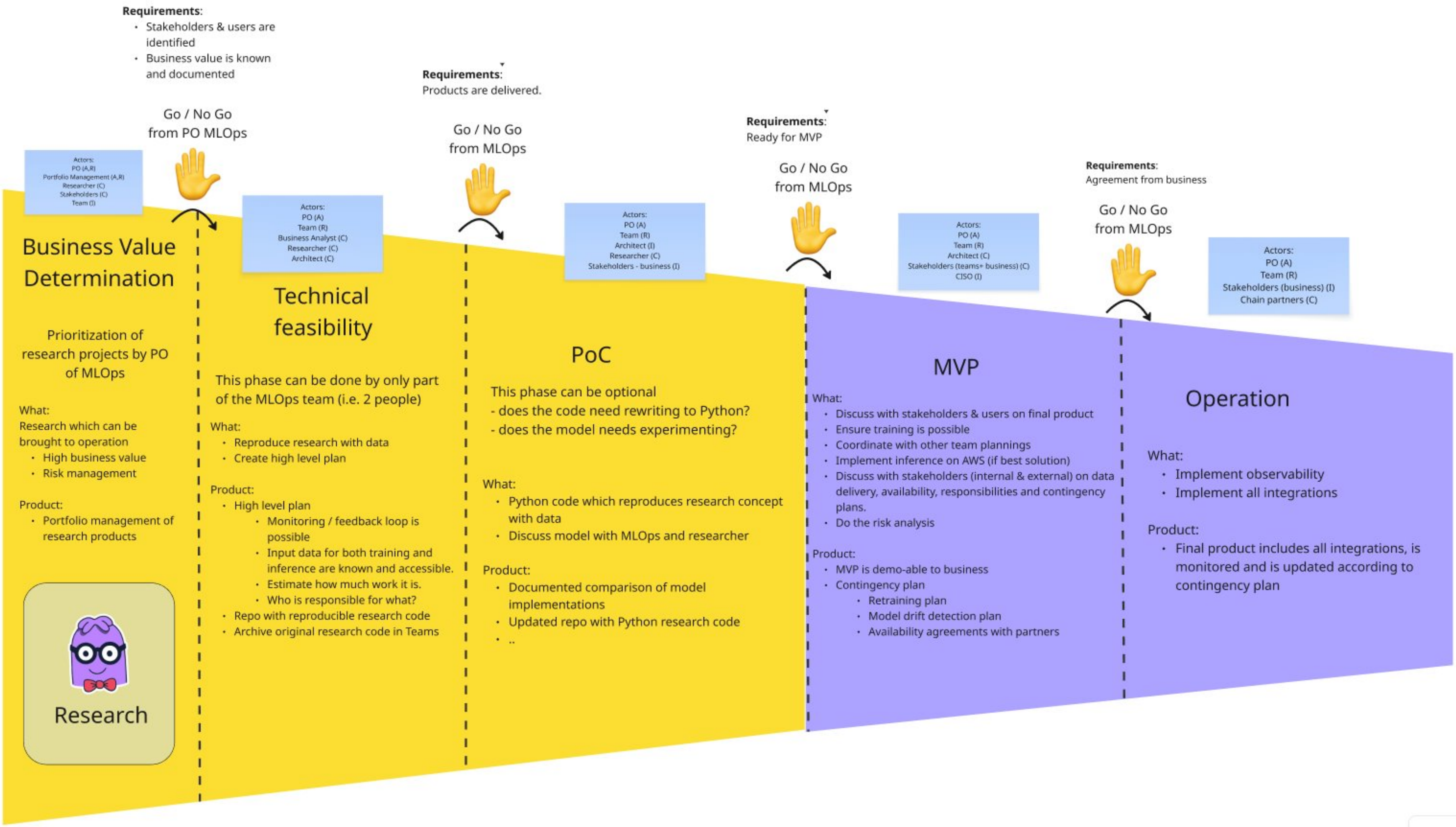
- Towards real-time operational service forecast using ECMWF inputs
- Operational testing and validation with users (KNMI weather room and NIPV)
- Extend the approach to other applications
- Build embedded MLOps culture at KNMI

Backup slides

AWS infrastructure Diagram



Project phases



Python model choice

- Several implementations:
  - Octave to python
  - Scikit-learn (RandomForestClassifier, binary classification, QuantileRegressor, HistGradientBoostingRegressor)
  - Ranger
  - Quantile Forest
  - **XGBoost (quantile regression)**

- Criteria:
  - License and maintenance of library
  - Performance metrics such as: RMSE, skill, pinball losses, prediction interval coverage and width
  - Integration with other frameworks (e.g. shap for explainability)
  - Ease of use and training time
  - Visualizations of forecast

Explainability for the weather room