

# eXplainable AI in regulated financial services

A contribution by Firamis GmbH, Consortium Partner

Dr. Jochen Papenbrock




# The project platform

## Existing use cases:



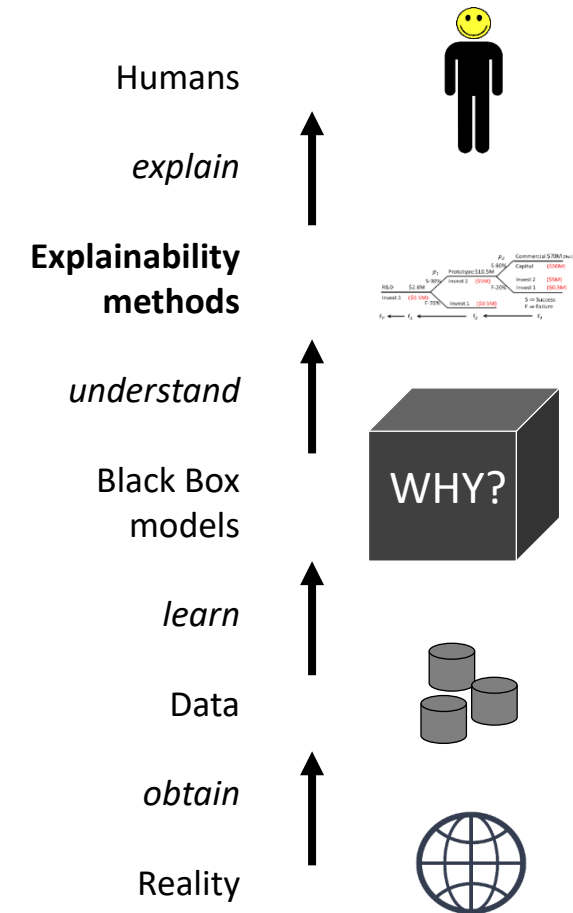
## Development:

-  is currently building a demonstration platform that wraps up the uses cases and substantially extends them by the latest developments in eXplainable AI

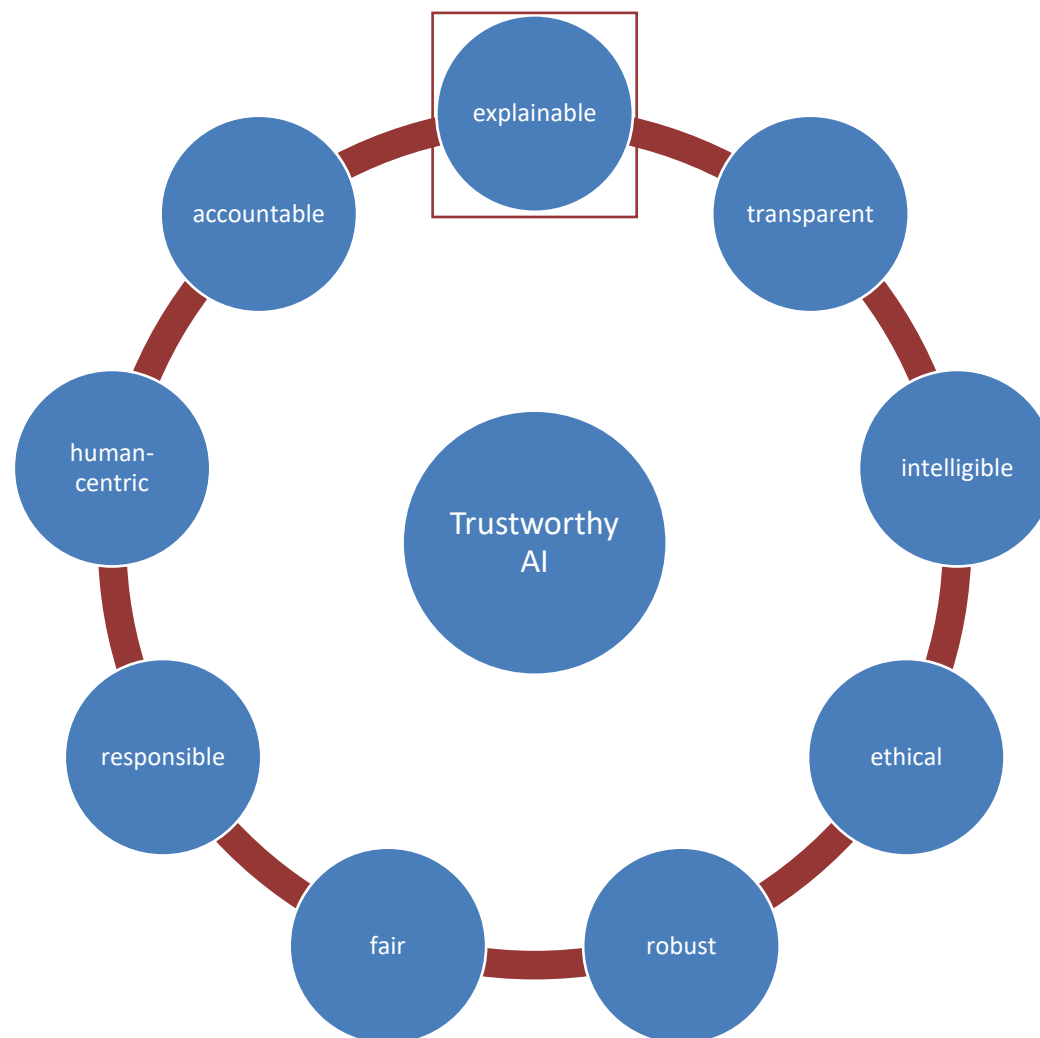
# eXplainable AI (XAI) – basics

Black box AI is not suitable in regulated financial services.

- XAI provides explanation approaches/technologies that translate AI models and data to a language understandable by humans
- Given an audience, an explainable Artificial Intelligence is one that produces details or reasons to make its functioning clear or easy to understand.
- When is a model explained? When you can't ask 'why' any more.
- By understanding the reasoning we can validate the correctness of the process.



# Components of trustworthy AI



High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence

3 out of 7 key requirements are linked to XAI.

# XAI - Importance

Both economic and regulatory need:

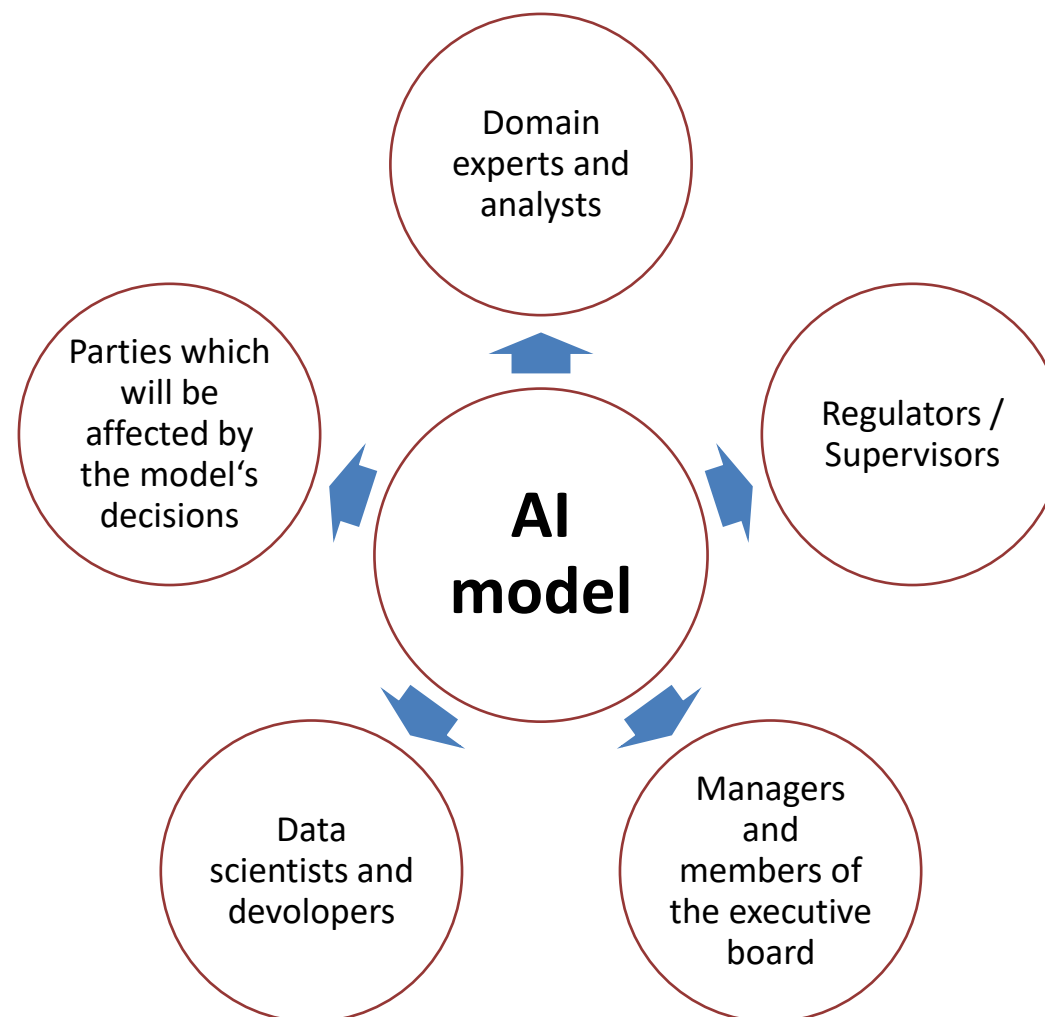
"The ability to explain model outputs to stakeholders is a major lever in ensuring compliance with expanding regulatory and public expectations and in fostering trust to accelerate adoption." \*)

XAI is an important building block in data-driven financial services.

\*) McKinsey: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/leading-your-organization-to-responsible-ai>

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825215. All material presented here reflects only the authors' view. The European Commission is not responsible for any use that may be made of the information it contains.

# XAI for which audience?



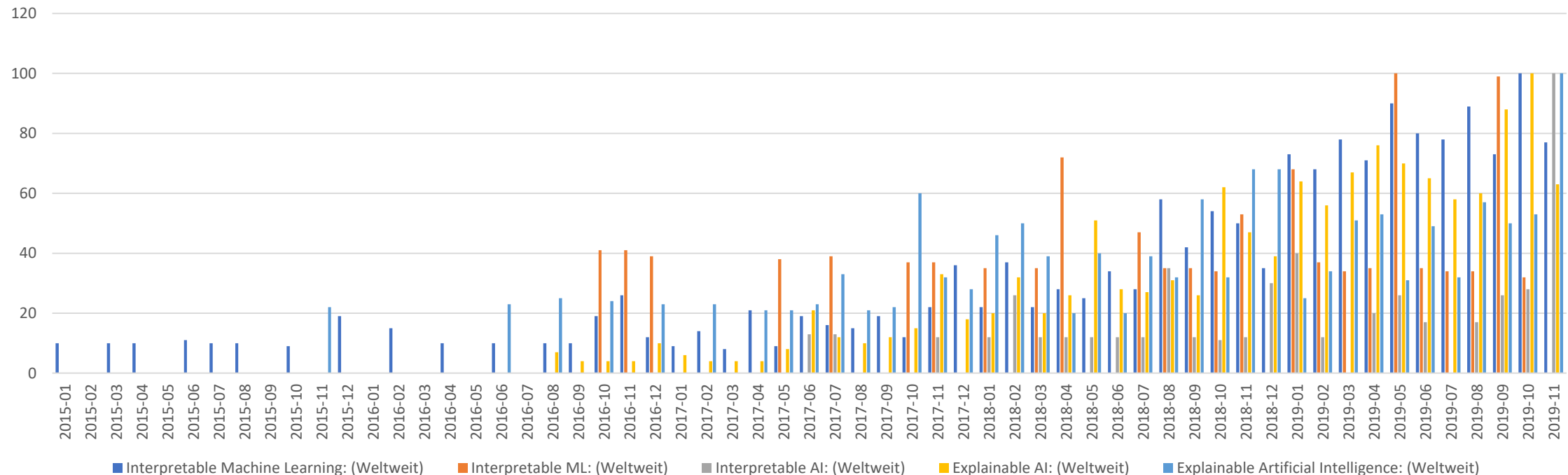
# XAI gets global attention

Many financial institutions and startups have started to embrace XAI.

Tech companies like Microsoft and IBM have launched XAI initiatives and programs.

XAI could emerge as industry standard

Google Trends



# eXplainable AI in regulated financial services

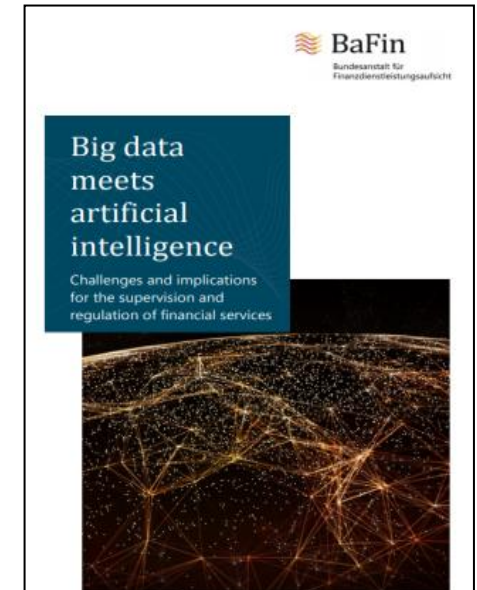
## *Challenges*



# Regulators

Regulators / supervisors are mentioning that Black Box ML models will no longer be accepted in the future

- Bank of England:
  - “‘Interpretability’ will be the focus - generally taken to mean that an interested stakeholder can comprehend the main drivers of a model-driven decision.”
- FSB: “lack of interpretability and auditability of AI and ML methods could become a macro-level risk”
- Financial Conduct Authority (FCA):
  - “In some cases, the law itself may dictate a degree of explainability.”



# Challenges

Survey over 106 financial firms, including banks, credit brokers, e-money institutions, financial market infrastructure firms, investment managers, insurers, non-bank lenders and principal trading firms

- ML models are often complex, large, non-linear and non-parametric
  - Difficult to understand their properties
  - Difficult to validate them
- ML methods are difficult to interpret
- Increased complexity makes model validation harder, Risks could be caused by a lack of ML model explainability
- ML use can alter nature, scale and complexity of IT applications and thus, a firm's IT risks
- Model validation and governance need to keep pace with ML developments.



BANK OF ENGLAND



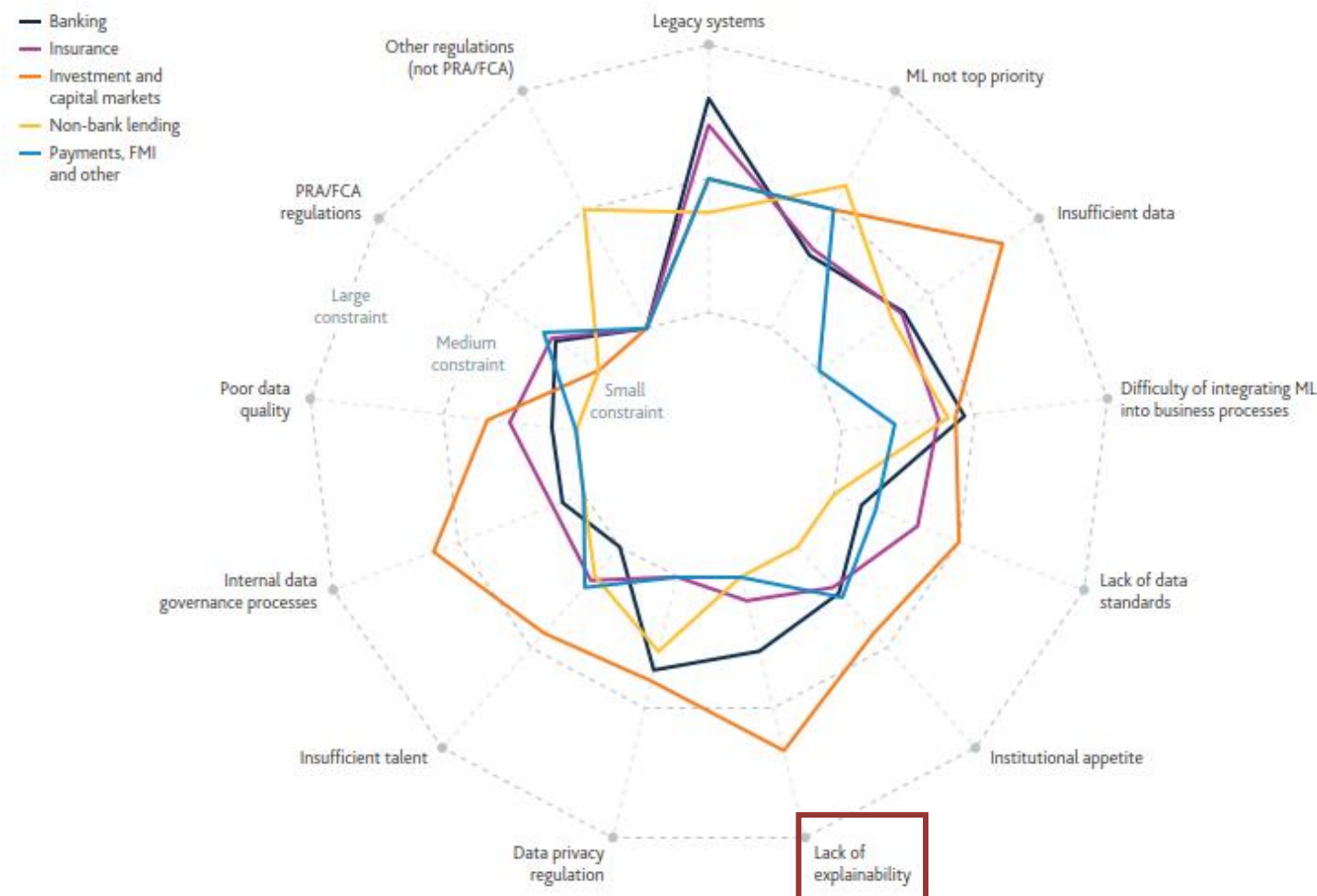
Machine learning in  
UK financial services

October 2019



# Challenges

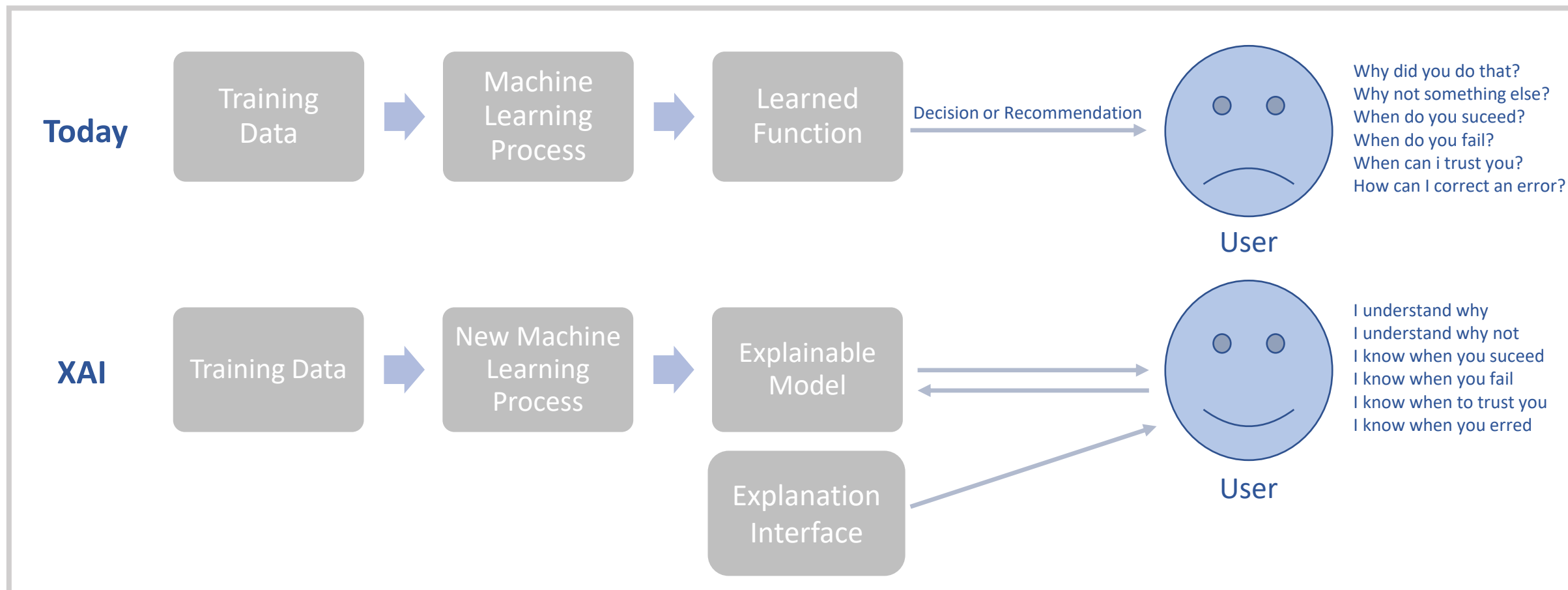
Several Constraints to the deployment of machine learning are mostly internal to firms:



# eXplainable AI in regulated financial services

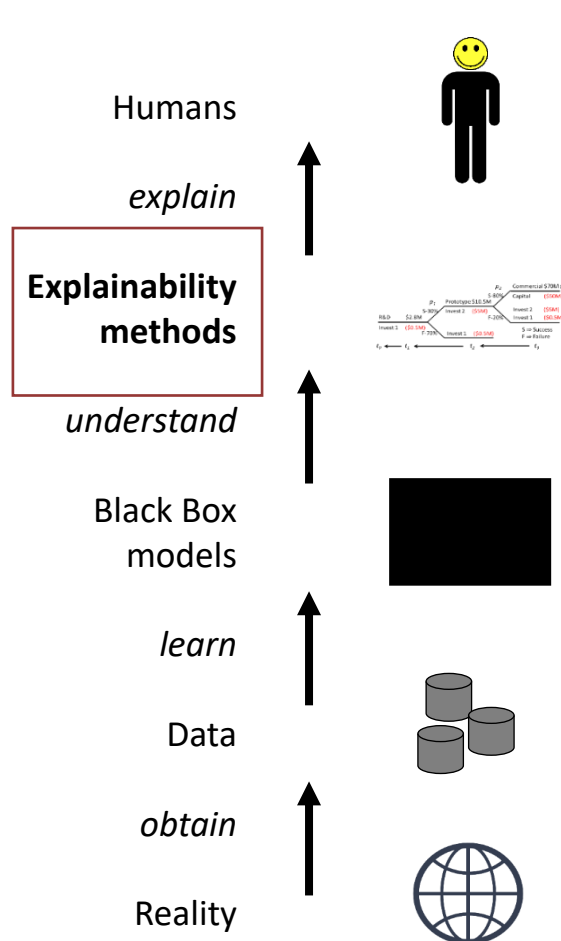
## *Solution*

# eXplainable AI – in simple pictures

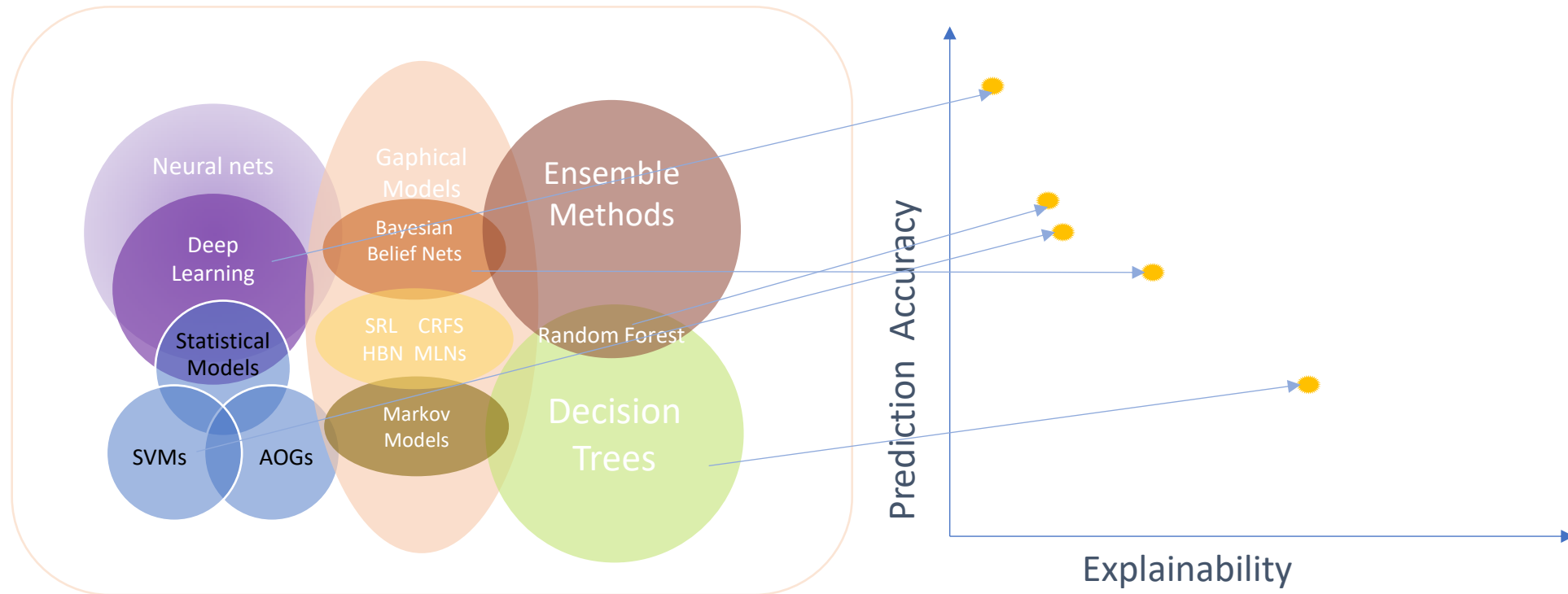


# How humans interface with XAI

- Visualisation
- Simplification
- Numbers / figures
- Examples
- Local + global features
- Feature interaction and sensitivity
- Textual



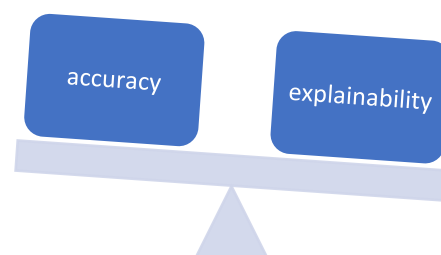
# Trade-OFF in complex ML approaches



# Trade-OFF with eXplainable AI

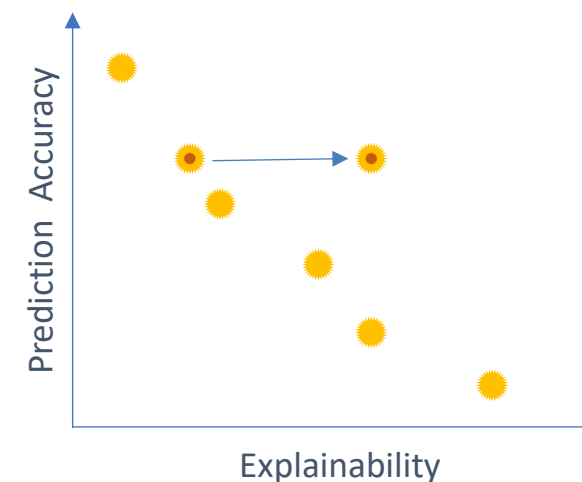
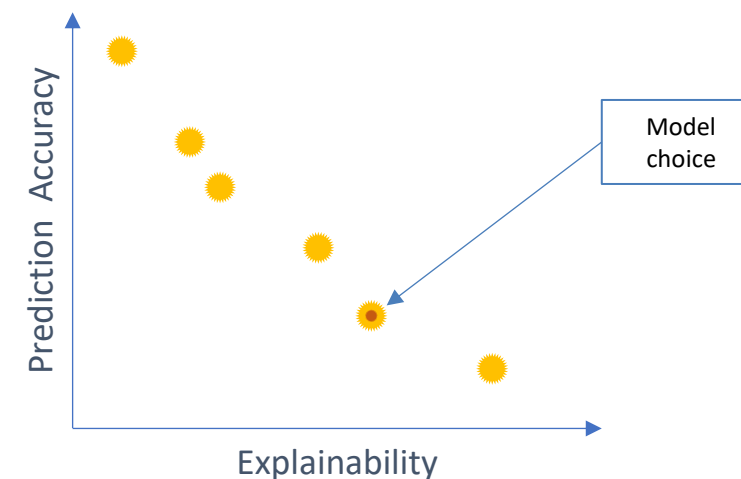
## TRADE-OFF SITUATION:

Strategic decision is made concerning fixed level



OR: start with a highly accurate model and introduce a layer of explanation technology

Does the output meet strategic requirements?





# Approaches to XAI

