



Project Profile

EXPLAIN

An explainable, interactive approach to machine learning

The AI Call 2021 project **EXPLAIN (EXPLANatory interactive Artificial intelligence for INdustry)** will create an end-to-end machine learning (ML) lifecycle for industrial domain experts that provides human supervision on artificial intelligence (AI) systems and ML models. This will increase the acceptance of AI/ML in industrial settings.

Addressing the challenge

AI holds great potential in industry, including improved throughput and sustainability, yet only 15% of AI projects started in 2022 will be successful. AI adoption rates are low in industrial settings due to a lack of trust in ML models without clear or transparent inner reasoning and outputs. As these models would be responsible for process efficiency and safety, domain experts and end-users must be given greater insights in the form of an explainable, interactive end-to-end ML lifecycle.

Proposed solutions

On a use-case basis, EXPLAIN aims to involve the right stakeholders early in the ML development process by enriching the conventional ML lifecycle with human-centred explanations. In the explanatory training phase, ML experts and domain experts will directly interact with an ML model and receive output explanations. Feedback can also be given through which the ML model can quickly be improved. In the explanation review, ML solutions will be validated by enabling domain experts to gain insights into the internal reasoning of the trained model, ensuring that relevant concepts have been learned from the data provided and uncovering misleading biases in the training dataset. The output of the ML model is then explained to end-users, who can provide feedback and trigger incremental explanatory training. These steps will be enabled via a seamless ML operations (MLOps) architecture within the lifecycle, avoiding the decoupling of

ML developers and software developers currently hampering the development and operation of large-scale software systems utilising ML/AI.

high safety requirements), allowing them to optimise their processes with AI/ML for the first time. In the process industry, for instance, 80% of the USD 20 billion in annual losses is estimated to be preventable. The estimated market access of new products and services from the consortium are expected to boost a turnover of EUR 500 million by 2026; they will also be able to position themselves competitively in the emerging global XAI market, predicted to grow by 18.4% annually to USD 21 billion in 2030.

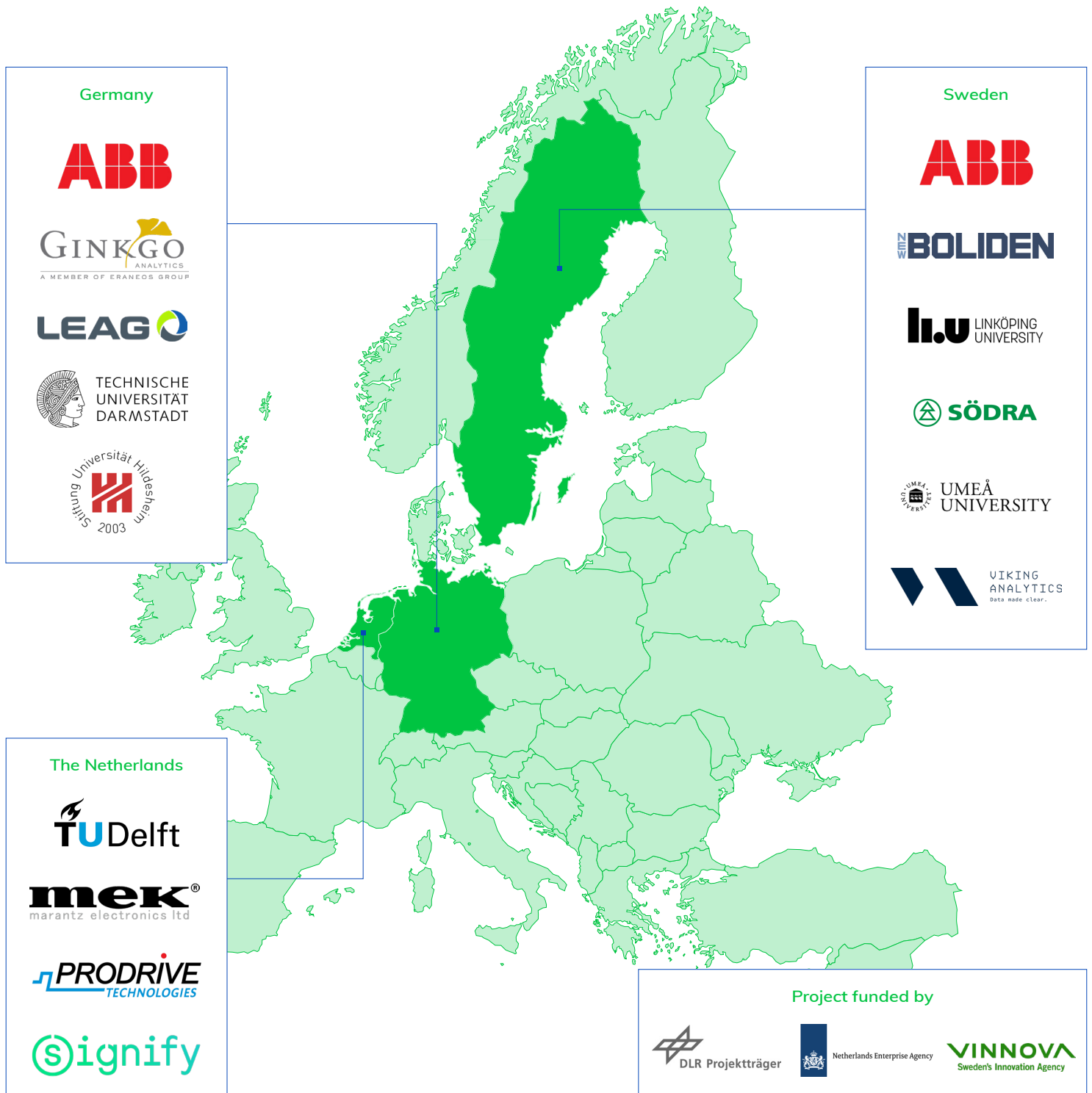


Control room operator troubleshooting an issue in a utility plant

Projected results and impact

EXPLAIN offers major research and business advancements in explainable AI (XAI). By introducing a more application/human-grounded evaluation to the XAI domain, the project creates insights on the usefulness of different explanation mechanisms and will help to increase the acceptance of ML models by domain experts. This will be especially useful for organisations for which explainability is a prerequisite (such as in industrial certification or critical industries with

In the longer term, the EXPLAIN approach could also be applied to other domains in which AI could supplement human decision-making (such as healthcare), giving the project an even larger reach.



Project start
May 2022

Project leader
Pablo Rodríguez Carrión, ABB

Project website
<https://explain-project.eu/>

Project end
May 2025

Project email
pablo.rodriguez@de.abb.com

ITEA is the Eureka RD&I Cluster on software innovation, enabling a large international community of large industry, SMEs, start-ups, academia and customer organisations, to collaborate in funded projects that turn innovative ideas into new businesses, jobs, economic growth and benefits for society. ITEA is part of the Eureka Clusters Programme (ECP).

<https://itea4.org>