

SIEMENS ADVANTA CONSULTING

Artificial Intelligence & Data Science in Practice

Justus Huemmer & Julian Spall | 11th of March 2025



Justus

Senior Consultant



Julian

Senior Consultant

**Back-
ground**

Electrical Engineering,
Information Technology

Mechanical Engineering,
Aerospace Engineering

Expertise

Data Analytics, Industrial IoT,
Assurance AI

Data Analytics, Industrial IoT,
Assurance AI



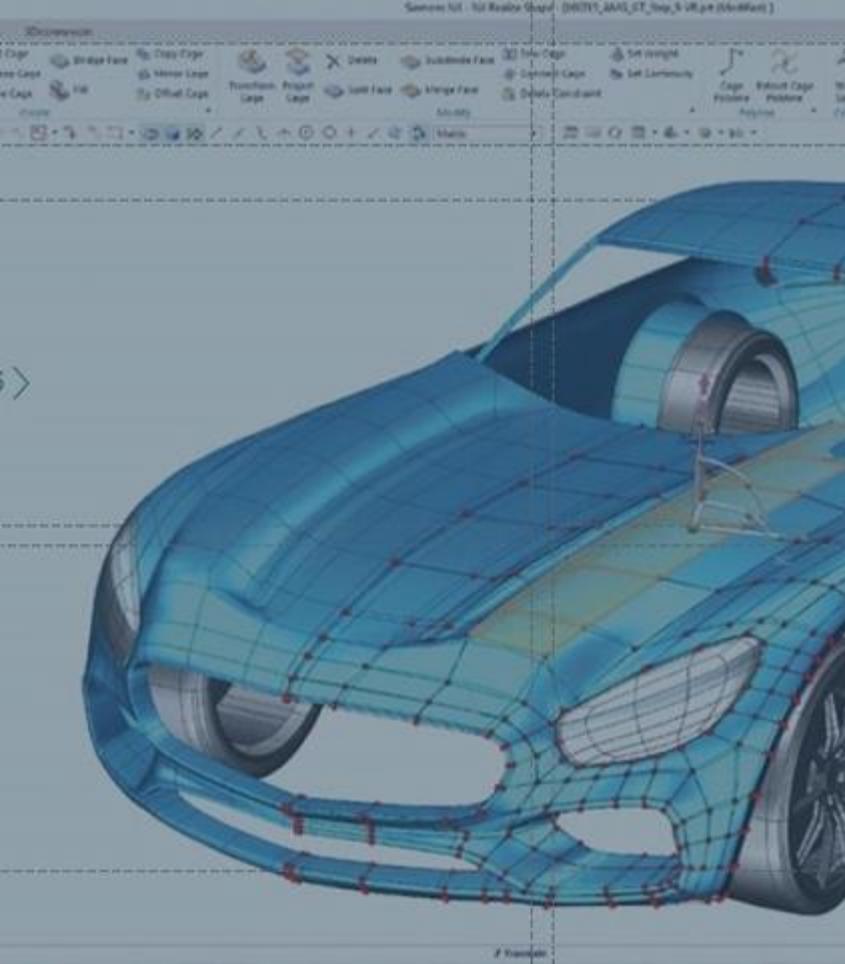
WHAT COMES
TO MIND WHEN
YOU THINK OF
SIEMENS?

Go to www.menti.com

and use the code

7425 7545





We commute in cars,
designed with **Siemens
Software** ...

... built in **factories**
automated by Siemens
...

... and charged with
renewable and
decentral **Siemens
Smart Grid**.



We travel in **trains made by Siemens**
...
...

... or in airplanes created with the help
of **Siemens Technology**.



We rely on **life saving medicine** to be quickly available on the markets.....

... enabled due to the **Siemens Innovations.**

Setup of Siemens AG

Businesses

Digital
Industries



Smart
Infrastructure



Mobility



Siemens
Advanta



Portfolio
Companies



Siemens
Healthineers¹

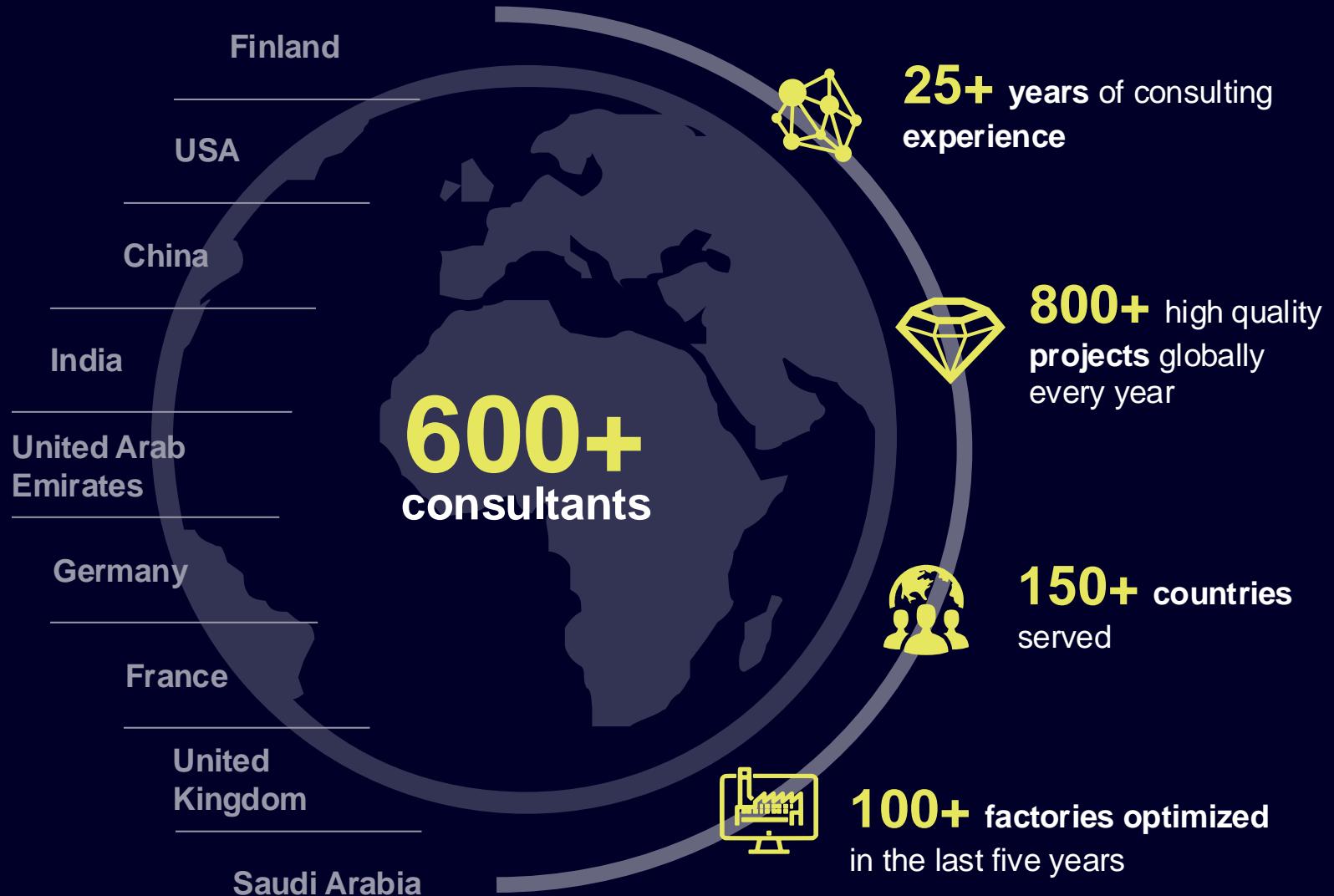


Countries

Service & Governance

¹ Publicly listed subsidiary of Siemens; Siemens' share in Siemens Healthineers: 75%

Siemens Advanta Consulting is a global consulting firm playing a key role in Siemens' digital journey



Our Siemens Advanta Consulting Portfolio



Our team is complementing the ADV Consulting strategy & digital portfolio with Data Science expertise – focusing on Enterprise & Industrial AI applications

Our Portfolio

Enterprise AI

harness artificial intelligence and automation to drive the **digital transformation** of **business processes**

Assurance AI

harness artificial intelligence to increase coverage, drive efficiency and **safeguard the enterprise**

Our Team

We employ **40+ data analytics consultants** in **3 countries**.



Our highly diverse team comes from **16 countries** and is fluent in **20 languages**.



Over the past two years, we have successfully **delivered >50** data science & AI **projects**.



Our **interdisciplinary** team has a strong background in academia and a broad data science and data engineering toolset.

Industrial AI

harness artificial intelligence to connect the physical and digital world and to drive the **digital transformation** of **industrial applications**

Cyber Security

harness artificial intelligence to protect IT networks, applications & data to **secure all business & manufacturing operations**

Our Data Analytics team



///

We will build AI into all our offerings based on a coherent data strategy.”

Dr. Roland Busch

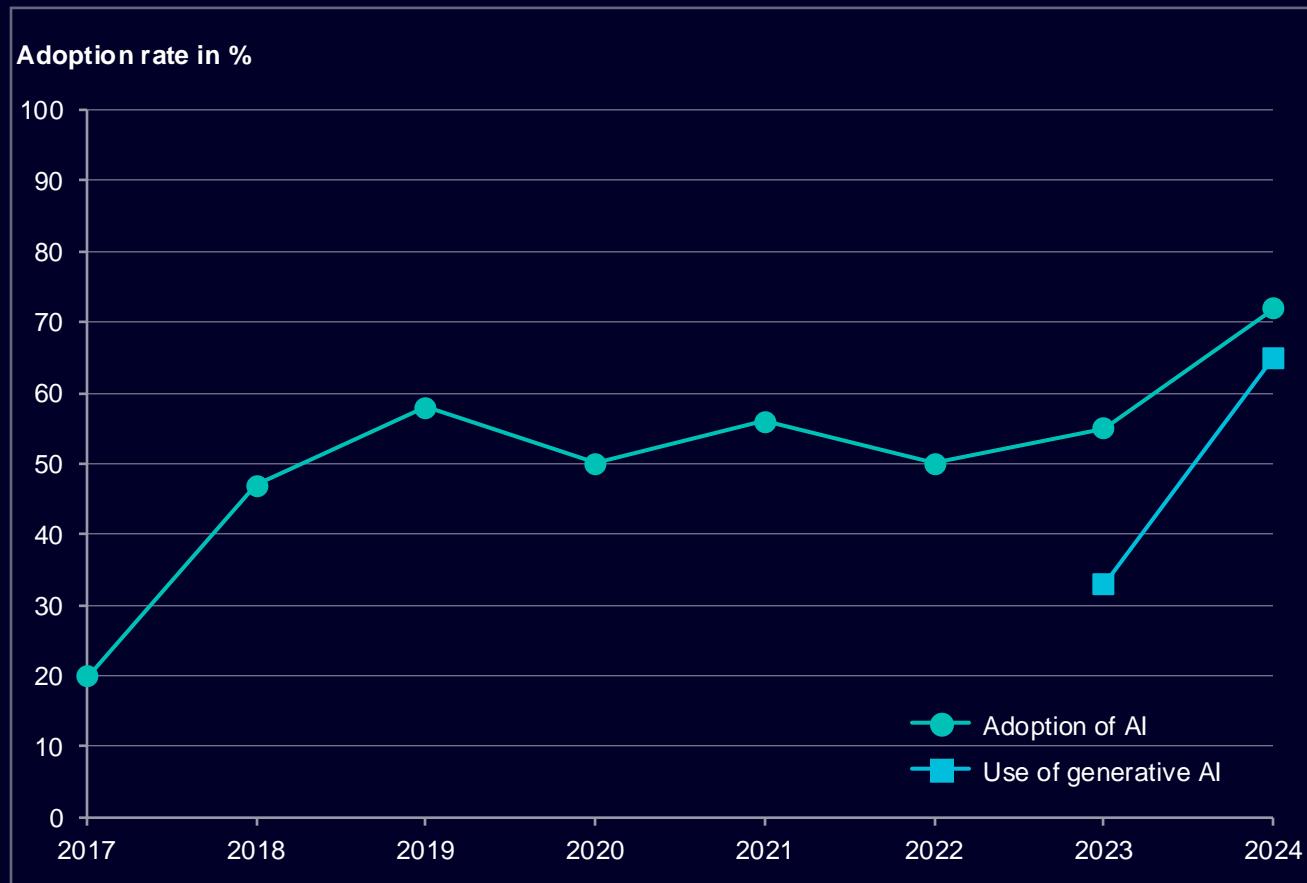
President and CEO of Siemens AG

Quote from Speech of the 2024 Annual Press Conference Munich, November 14, 2024.



The importance of AI – AI adoption worldwide has dramatically increased in the past year boosted by GenAI

Overview of organizations, that have adopted AI in at least one business function



The adoption of advanced AI technologies, particularly generative AI (GenAI) and large language models (LLMs), has accelerated significantly across various industries.



Global Adoption of Generative AI: As of May 2024, 65% of surveyed businesses worldwide reported regularly using generative AI, a substantial increase from the previous year.

Source: McKinsey Global Survey on AI, 1,363 participants at all levels of the organization, Feb 22 – Mar 5, 2024

The importance of AI – AI adoption worldwide has dramatically increased in the past year boosted by GenAI

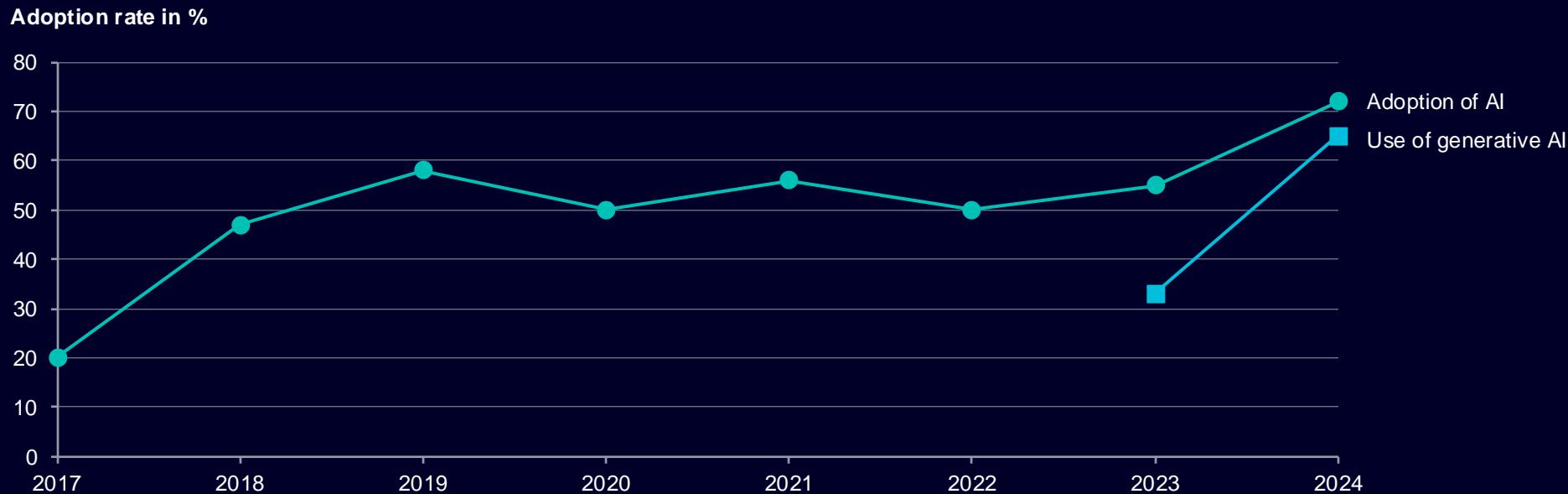


The adoption of advanced AI technologies, particularly generative AI (GenAI) and large language models (LLMs), has accelerated significantly across various industries.



Global Adoption of Generative AI: As of May 2024, 65% of surveyed businesses worldwide reported regularly using generative AI, a substantial increase from the previous year.

Overview of organizations, that gave adopted AI in at least 1 business function:

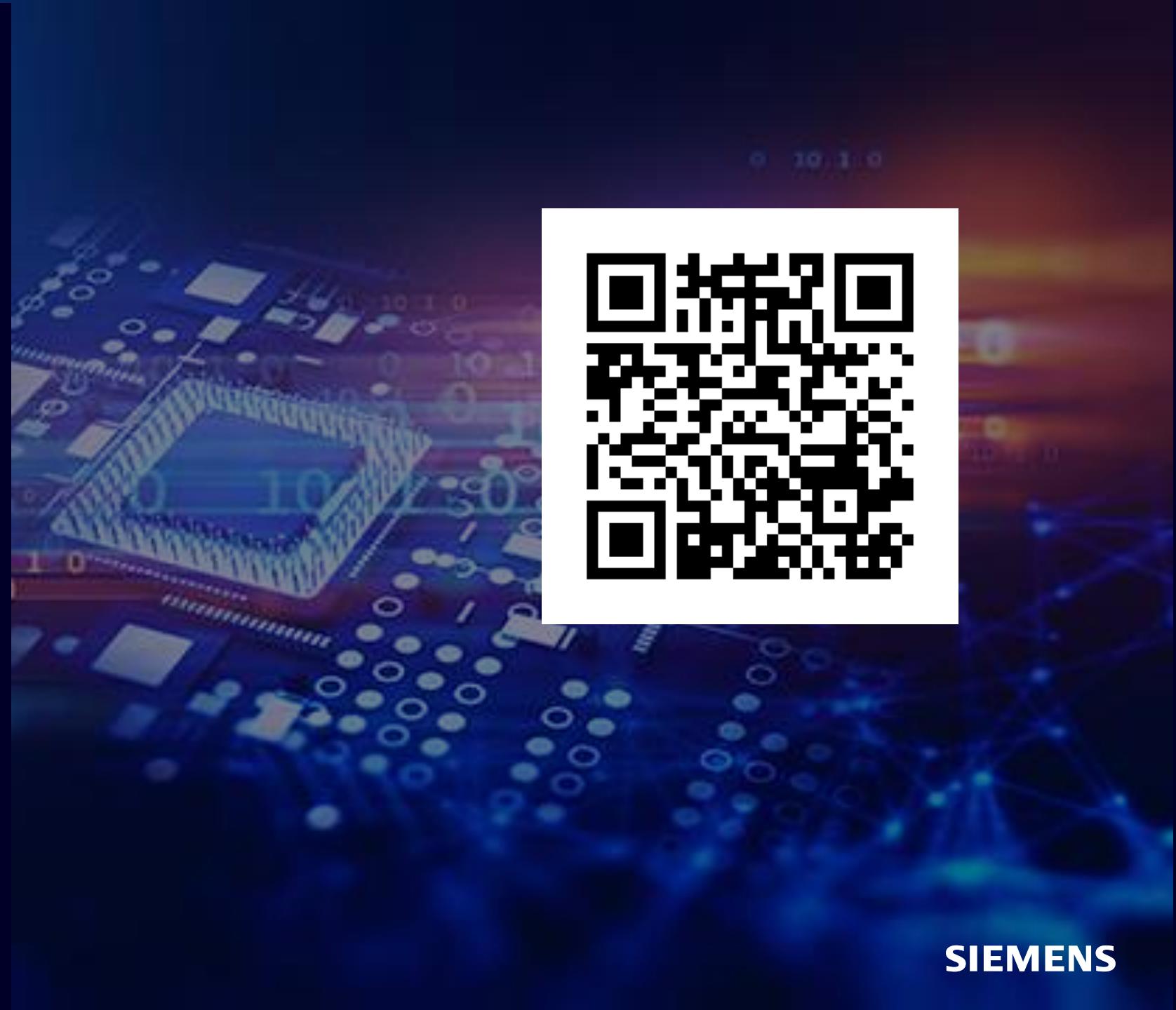


Source: McKinsey Global Survey on AI, 1,363 participants at all levels of the organization, Feb 22 – Mar 5, 2024

**What do you
estimate is the
percentage of Data
Science projects
that never go into
production?**

Go to www.menti.com
and use the code

7425 7545



What do you estimate is the percentage of Data Science projects that never go into production?



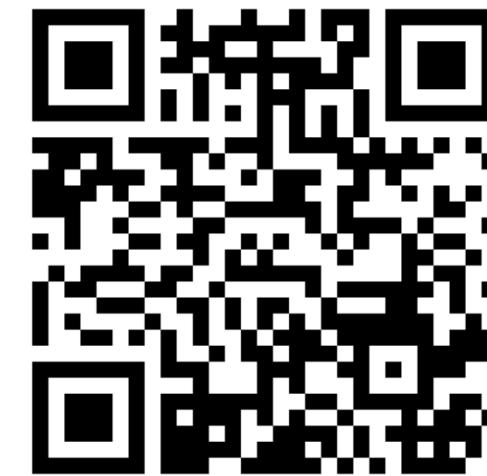
85%
FAIL

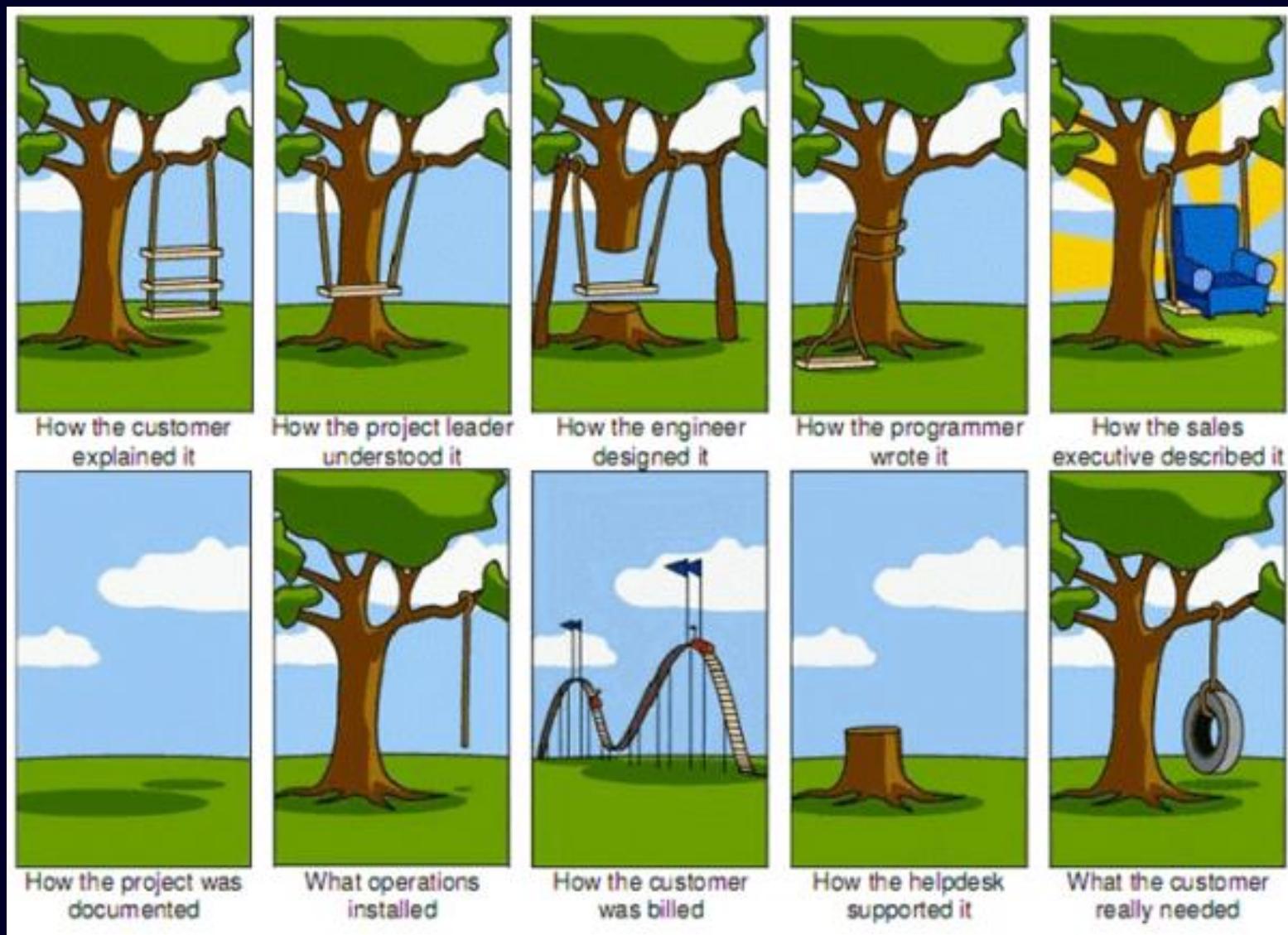
Source: Gartner 2017

**What do you think
are the reasons for
this high rate of
failure?**

Go to www.menti.com
and use the code

7425 7545





**Data Analytics projects fail
But what can you do about it?**



**Fail fast and make
sure
that you learn**



**Increase chances
of success**

Data Analytics projects fail But what can you do about it?

In the PoV phase, we want to address questions regarding the desirability, feasibility, and viability of the solution.

Question & Definition

Desirability	Feasibility	Viability
Desirability The solution is needed by the organization	Feasibility The solution must help the business professional effectively in time or quality	Viability The solution must generate a net positive value add for the organization.
Assess Reusability and scale • Can this input be interchangeable and be used in other businesses?	Model performance: • What is the most effective way to wanted result? • Can existing solutions and features improve the current development and outcome?	Business case evaluation: • What is the cost of the use case? • What is the expected value add of the user need?
Define the questions to be answered during the POC!	Implementation evaluation: • What is the estimated implication (efforts) for full development of the solution and its maintenance?	Data and infrastructure feasibility: • How does sensitivity or criticality of data affect the solution infrastructure setup?

Page 21 © Siemens Advanta Consulting 2025

For discussion

SIEMENS

Fail fast and make sure that you learn



There is more to a successful data analytics product than just a state-of-the-art algorithm

Machine Learning Algorithm

Open-source, freely available on the market, multi-purpose

Page 22 © Siemens Advanta Consulting 2025

SIEMENS

Increase chances of success



In the PoV phase, we want to address questions regarding the desirability, feasibility, and viability of the solution

Question & Definition

Aspect	Desirability	Feasibility	Viability
Define the questions to be answered during the POC!	<p>The solution is needed by the organization</p> <p>Reusability and scale</p> <ul style="list-style-type: none"> • Can the input/tool be interchangeable and be used in other Businesses? <p>Integration into the business process</p> <ul style="list-style-type: none"> • Can the solution be integrated into the business tools and system landscape? 	<p>The solution must help the business professional effectively in time or quality</p> <p>Model performance:</p> <ul style="list-style-type: none"> • What is the most effective way to wanted result? • Can existing solutions and features improve the project development and outcome? • Can we reach satisfactory results in terms of defined metrics (e.g. accuracy?) <ul style="list-style-type: none"> • Who evaluates the quality? • What is the evaluation set? • How many examples are needed to make sure that the solution is validated? <p>Data and infrastructure feasibility</p> <ul style="list-style-type: none"> • How das sensitivity or criticality of data affect the solution infrastructure setup? 	<p>The solution must generate a net positive value add for the organization.</p> <p>Business case evaluation</p> <ul style="list-style-type: none"> • What is the cost of the use case? • What is the expected value add of the use case? <p>Implementation evaluation</p> <ul style="list-style-type: none"> • What is the estimated implication (efforts) for full development of the solution and its maintenance?

Technical POC

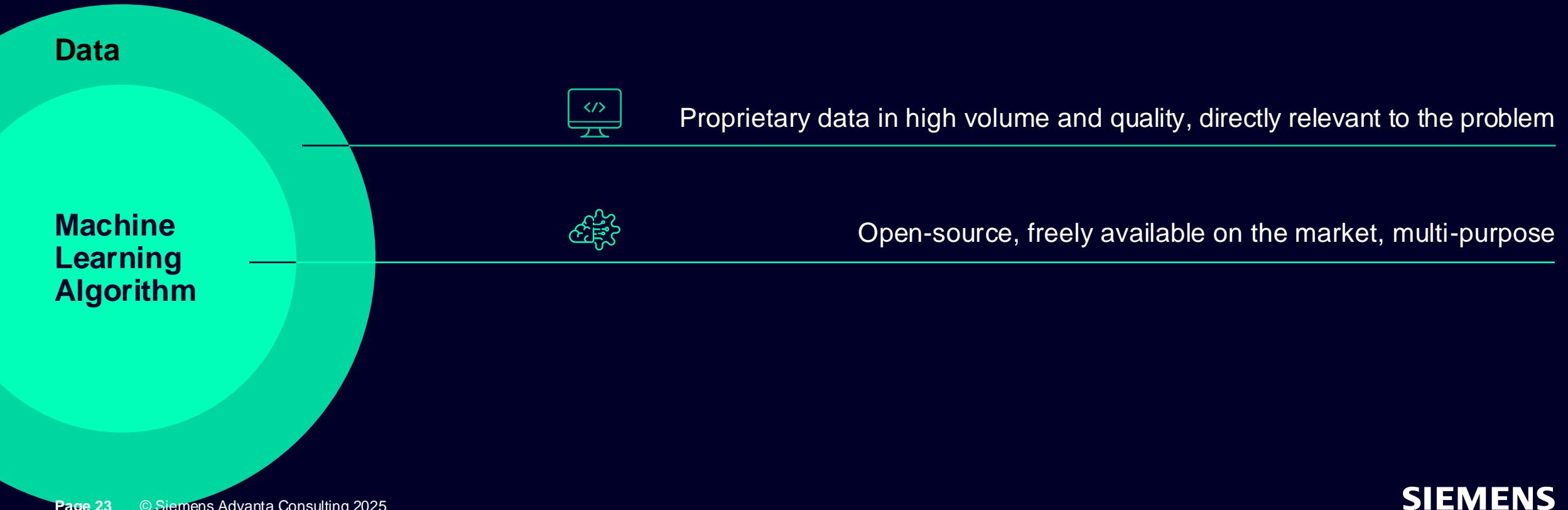
There is more to a successful data analytics product than just a state-of-the-art algorithm

**Machine
Learning
Algorithm**

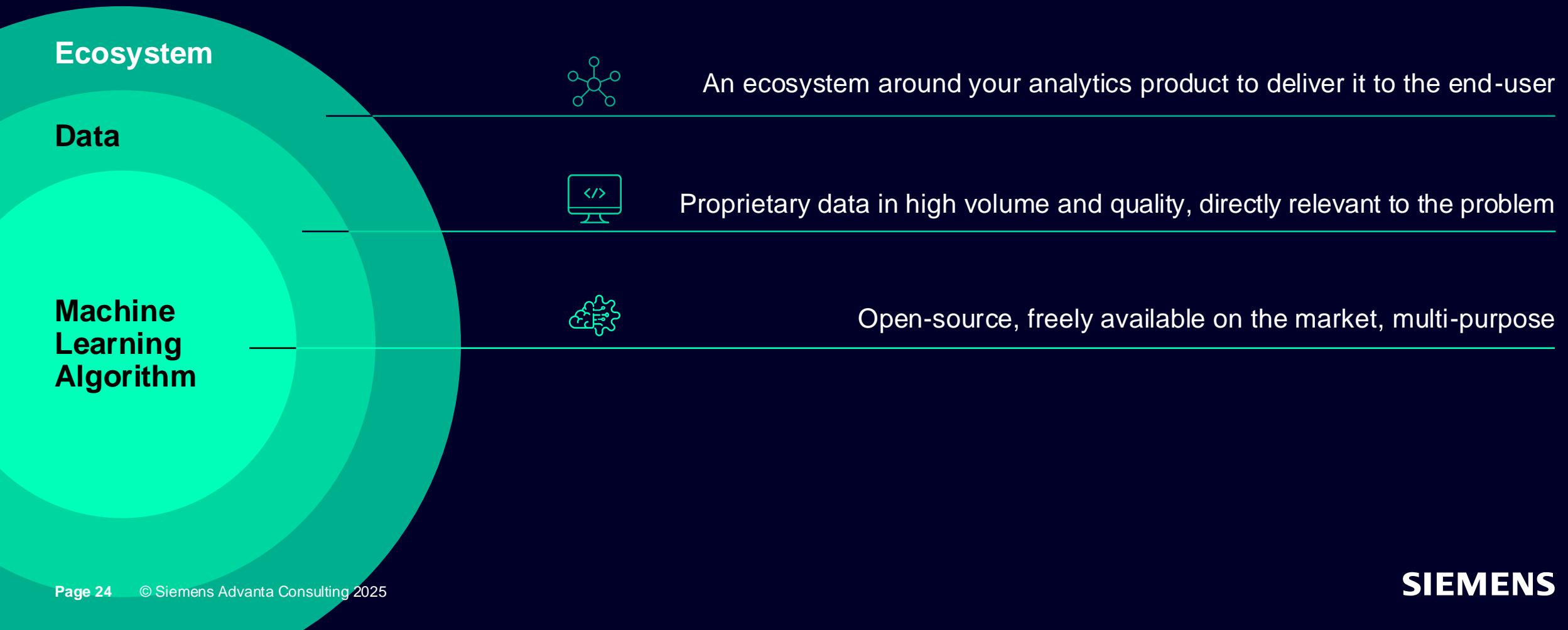


Open-source, freely available on the market, multi-purpose

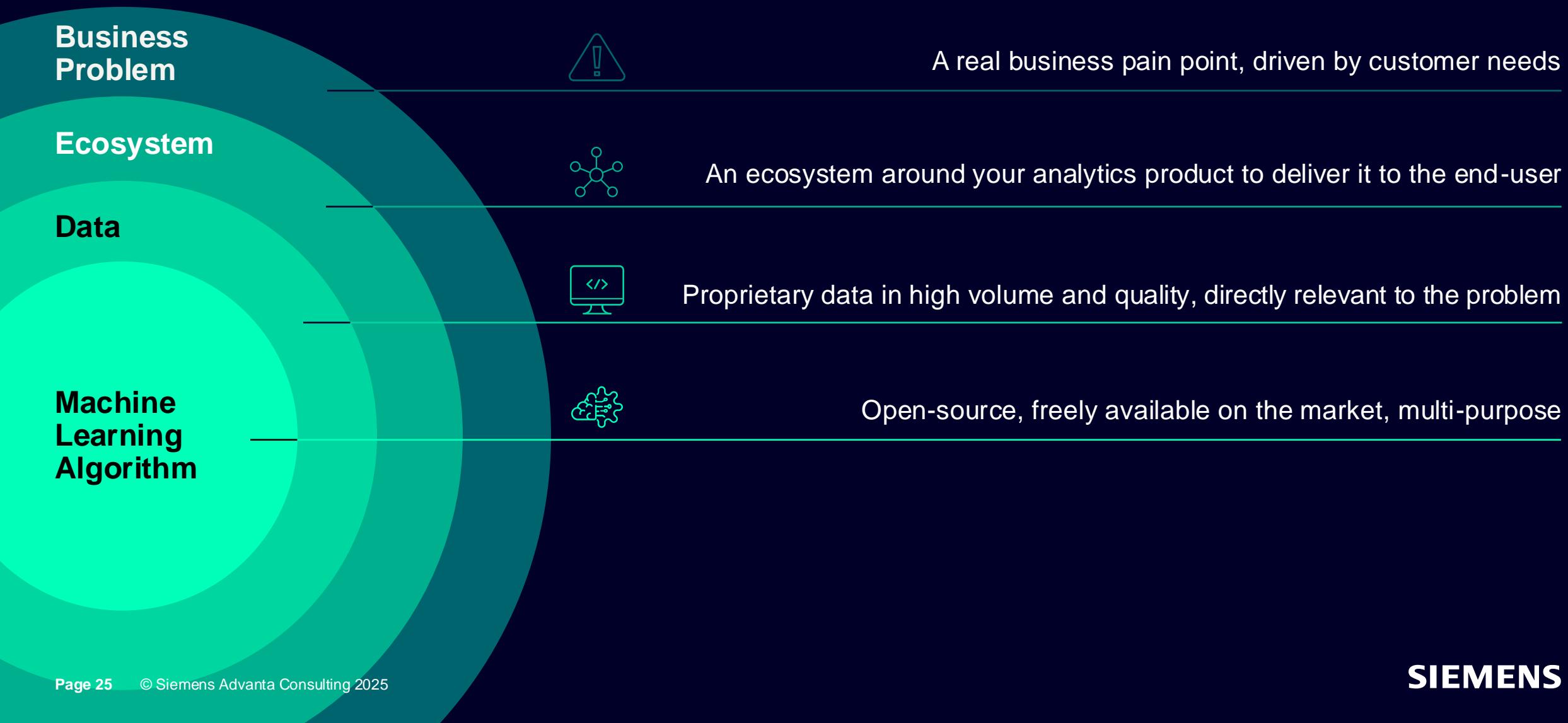
There is more to a successful data analytics product than just a state-of-the-art algorithm



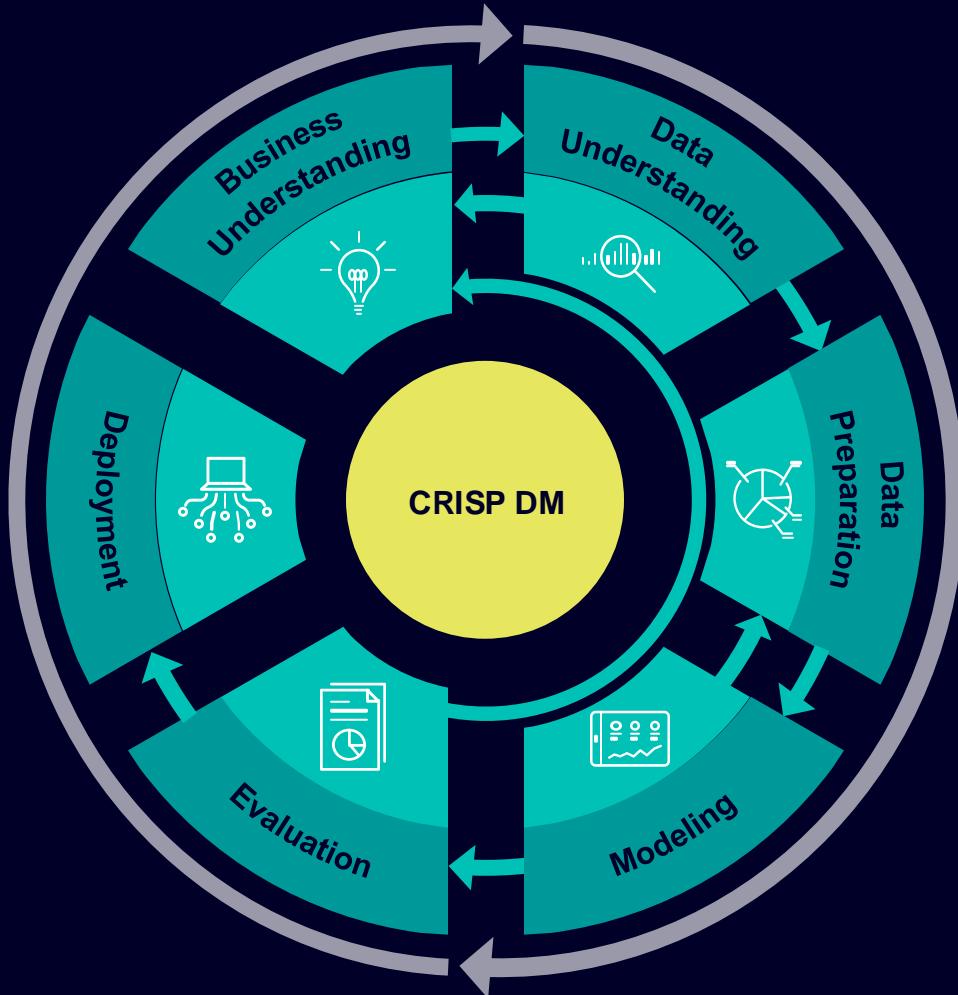
There is more to a successful data analytics product than just a state-of-the-art algorithm



There is more to a successful data analytics product than just a state-of-the-art algorithm

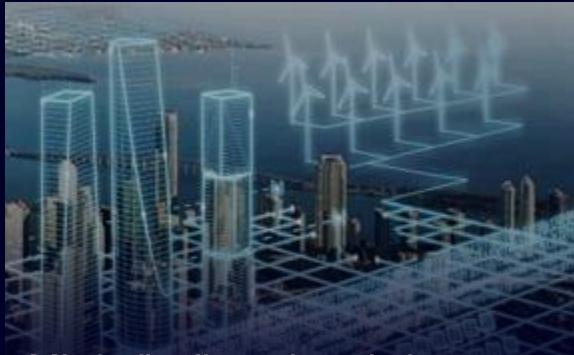


The reasons for Data Science projects to fail are manifold – the Cross Industry Standard Process for Data Mining (CRISP-DM) is one lever to mitigate the risk of failure



Source: CRISP-DM 1.0 (Chapman et al.)

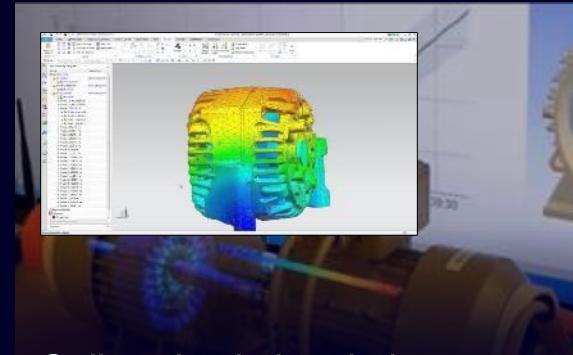
Time to put AI into work - towards the use of Artificial Intelligence in a comprehensive context



ML & distributed analytics –
Intelligent grid controller



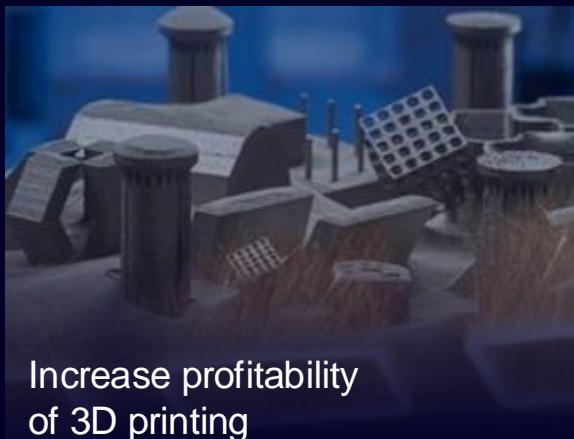
Teaching trams to drive
autonomously



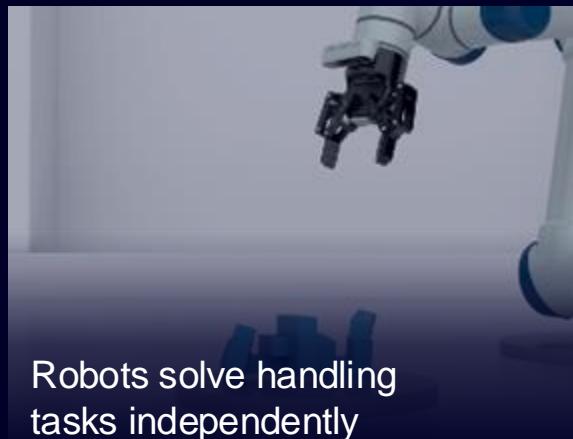
Online simulation during
operation



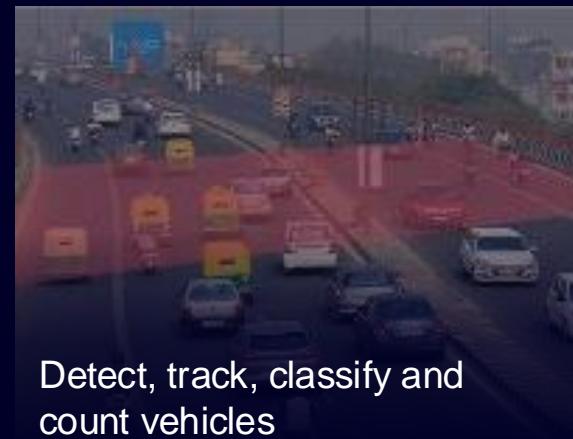
Optimize operation of gas
turbines (<15 – 20% NOx)



Increase profitability
of 3D printing



Robots solve handling
tasks independently



Detect, track, classify and
count vehicles



Optimize processing
time and quality

The framework is industry agnostic and has helped us in deploying data science solutions across a wide range of industries & customers

Cloud service auditor

Detection of sensitive personal data on Corporate SharePoint to reduce data privacy risks



Food and beverage industry

Digital solutions to enhance visibility, efficiency and decision-making in manufacturing plants



The framework is industry agnostic and has helped us in deploying data science solutions across a wide range of industries & customers

Cloud service auditor

Detection of sensitive personal data on Corporate SharePoint
to reduce data privacy risks



Food and beverage industry

Digital solutions to enhance visibility, efficiency and decision-making
in manufacturing plants



GenAI-Powered Data Breaches Detection



CLIENT

Finance and assurance team



CLIENT CHALLENGES

A multinational corporation with **300+ subsidiaries** faced a **critical data privacy risk** due to sensitive employee information - such as **passport details, Social Security numbers, bank statements, and salary records**—being stored on a **corporate SharePoint with insufficient access restrictions**.

- **Risk of Data Breaches:** Confidential data was stored in locations with less restrictive access controls, increasing vulnerability to potential cyber threats or unauthorized access.
- **Manual & Time-Consuming Audits:** The audit team spent significant effort manually screening and verifying files for sensitive information.
- **Non-Reusable & Non-Scalable Solutions:** Existing processes were inefficient, lacked automation, and could not scale across large datasets.



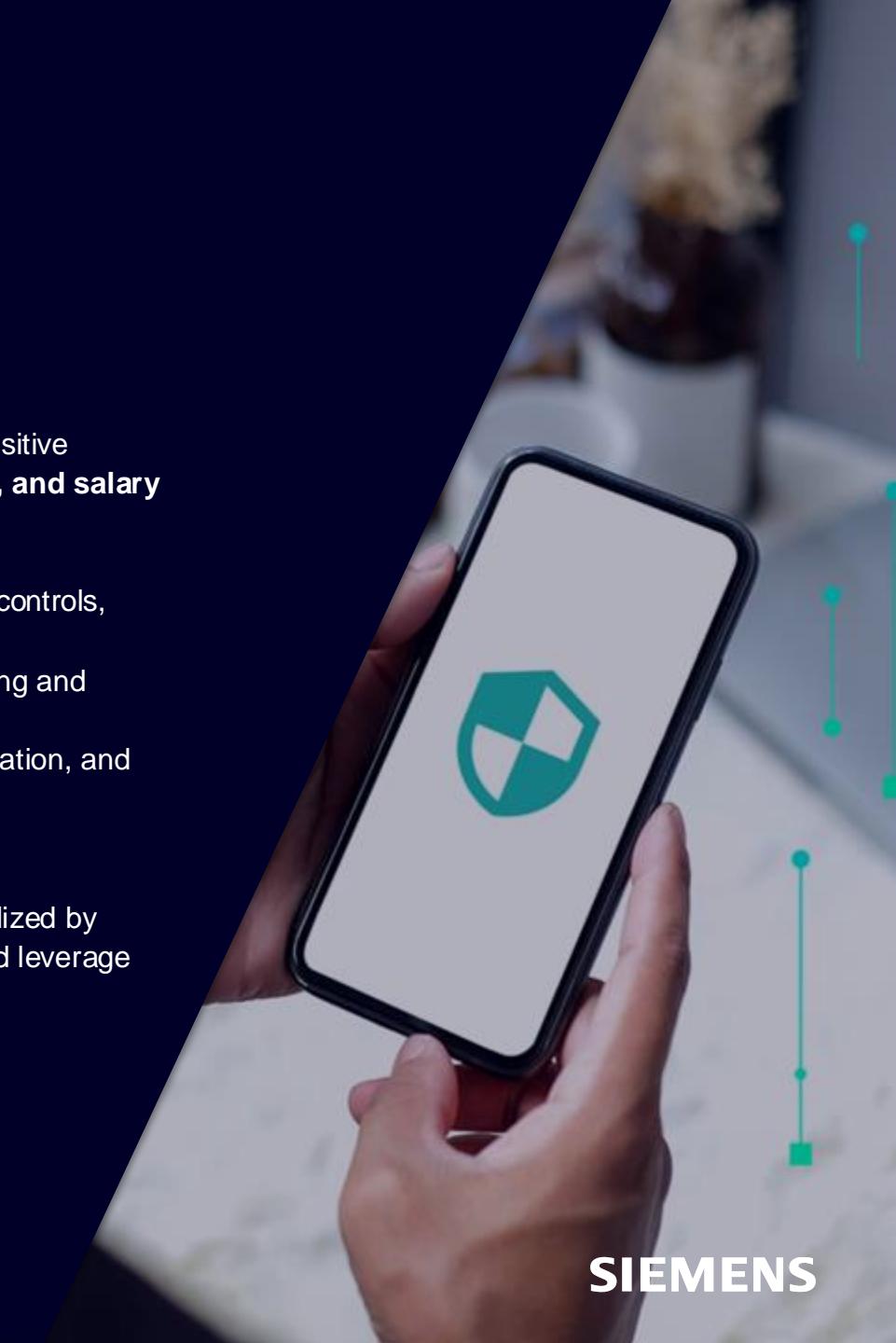
OUR SOLUTION

An application with web crawling functions **enhanced with Generative AI** can be seamlessly utilized by internal users to **search and identify files**, automatically **screen for sensitive information**, and leverage **GenAI-powered intelligence** to improve the accuracy and efficiency of file detection.



CLIENT BENEFITS

- ✓ Time saving & Increased efficiency
- ✓ Cost reduction
- ✓ Improved accuracy
- ✓ Scalable solution





Early Engagement & Collaboration: Building a clear & aligned project vision

End User



Challenges

- Risk of Data Breaches
- Manual & Time-Consuming Audits

- Non-Reusable & Non-Scalable Solutions

Approach

- Experience the current solutions firsthand
- **Deep Client Collaboration:** Work closely with stakeholders to identify pain points, map workflows, and uncover key challenges.
- **Actionable & Aligned Use Cases:** Develop practical, real-world use cases with the client to define the project scope and ensure the solution directly addresses business needs.

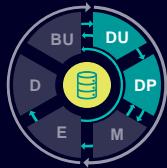


Lessons Learned

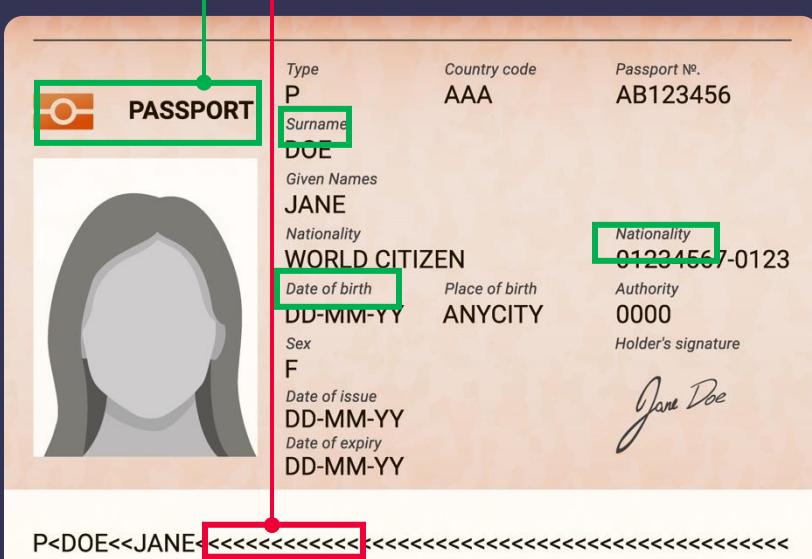
- **Early-stage discussions** on application features and scope promote clarity and alignment for the entire project
- Engaging end users **early** helps harmonizing the diverse needs and desires of all stakeholders



Data Analysis and Pattern Recognition: Effectively identifying sensitive information across various file types



Identifiable Patterns



Effective Keywords

- Passport
 - "PAN"
 - Social Security Number
 - XXX-XX-XXXX
 - Salary Information/salaries
 - Date of birth
 -

Approach

- Clear focus on effective identifiable patterns and keywords
 - Text extraction pipeline with use of OCR technology

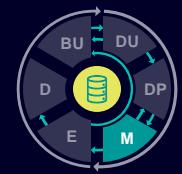
Lessons Learned

- Domain knowledge > Technology
 - Working with client **side by side** for seamless deployment and immediate feedback

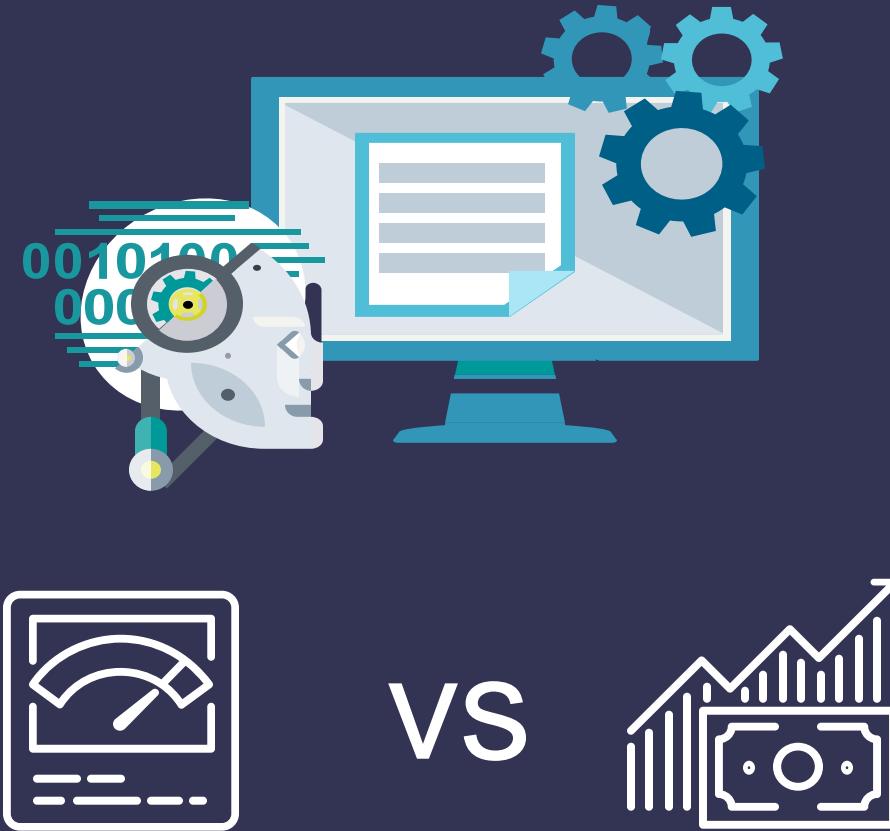
Challenges

- Precisely defining Personally Identifiable Information (PII) requires domain knowledge and a thorough understanding of the specific data types that qualify as PII





The Real Deal: Applying multiple techniques to minimize false positives and maximize detection accuracy



Approach

- Advanced NLP +
- Regex Searching +
- Advanced analytics package +
- GenAI

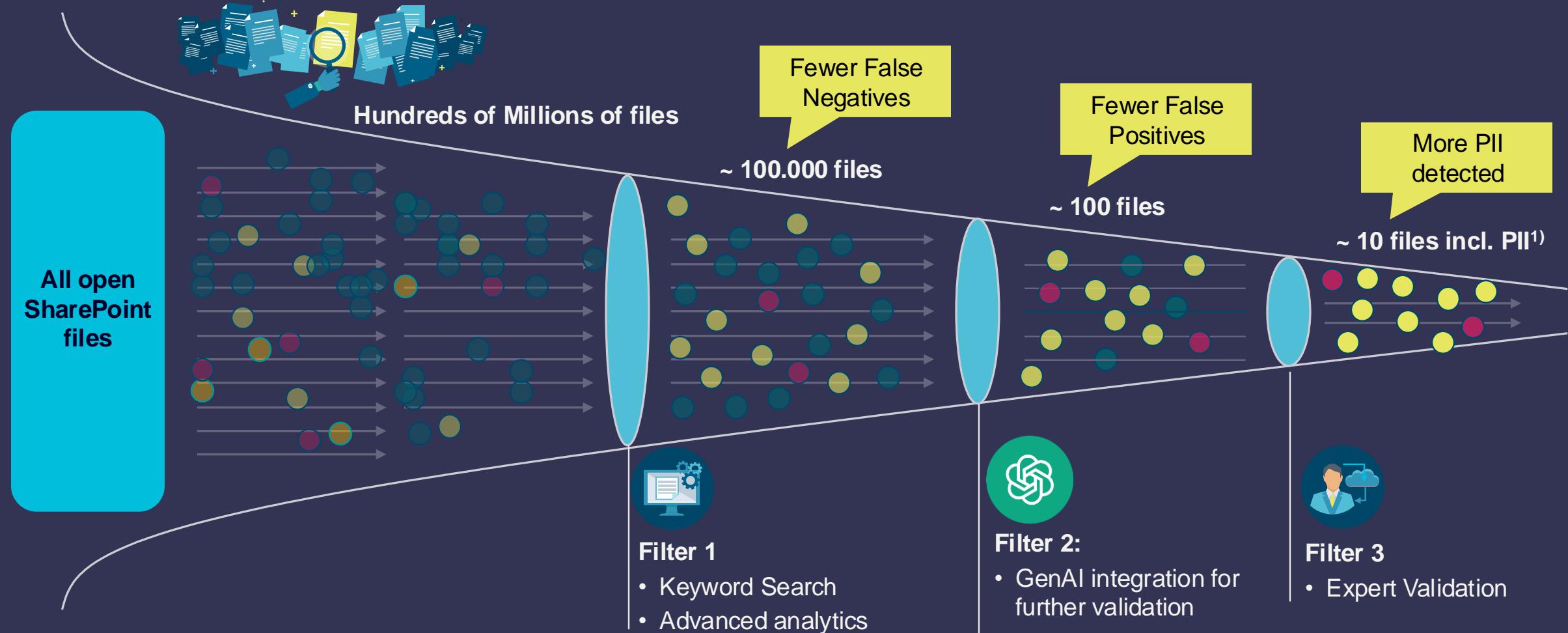
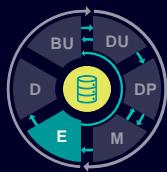


Lessons Learned

- Prioritization matters
 - Trade-off between performance (more advanced models) vs. time & cost
- Selection must be context-driven, prioritizing efficiency and business needs
- Communication is always the key

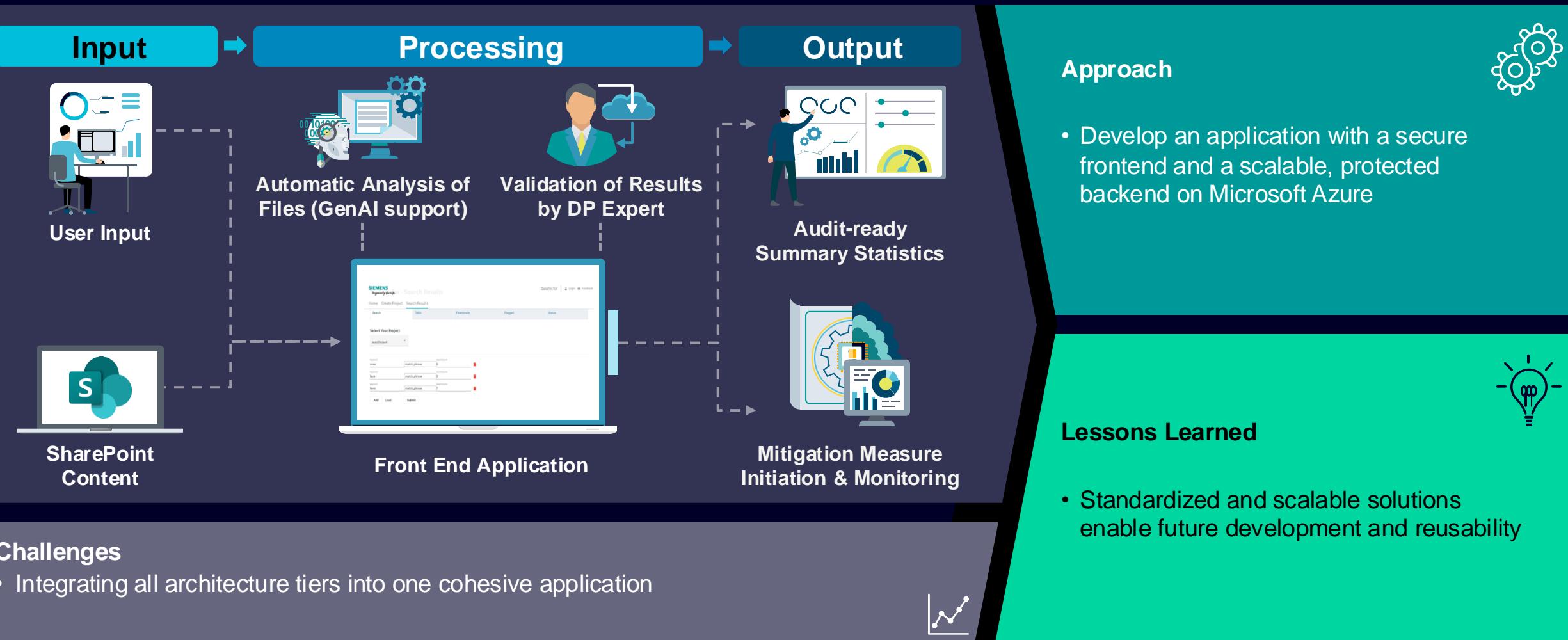


The Real Deal: Applying multiple techniques to minimize false positives and maximize detection accuracy





Developing on standardized infrastructure: The end-to-end application is engineered for scalability across various use cases and long-term maintainability



The framework is industry agnostic and has helped us in deploying Data Science solutions across a wide range of industries & customers

Cloud service auditor

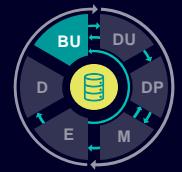
Detection of sensitive personal data on Corporate SharePoint to reduce data privacy risks



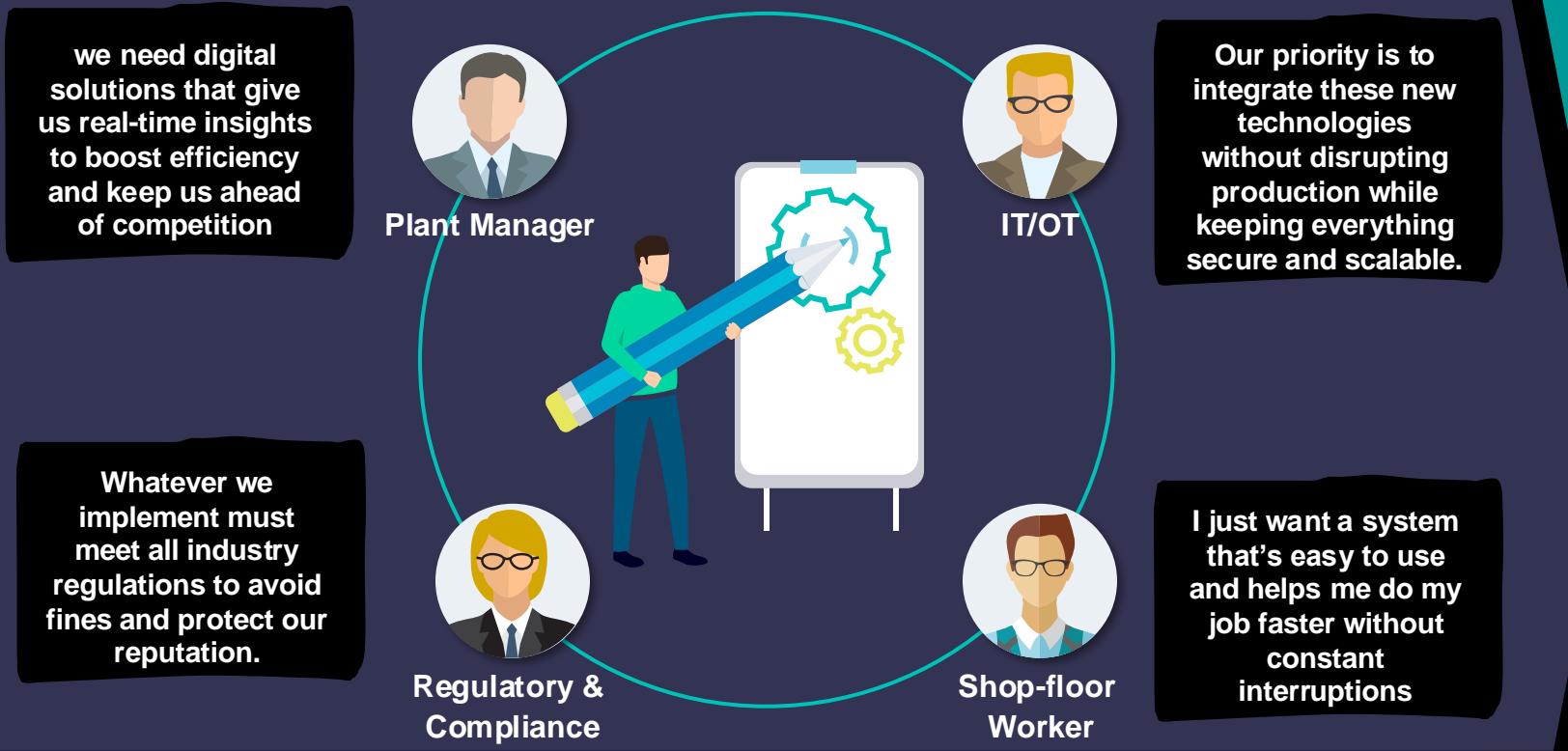
Food and beverage industry

Digital solutions to enhance visibility, efficiency and decision-making in manufacturing plants





Identifying application benefits and prioritizing use cases for application implementation via workshops



Challenges

- Discover the potential financial impacts and clearly define how each application can be effectively utilized in specific scenarios



Approach

- Conduct plant leaders' workshop that aims to identify, brainstorm, evaluate, and prioritize use cases for application implementations by addressing operational challenges and aligning on a strategic roadmap.



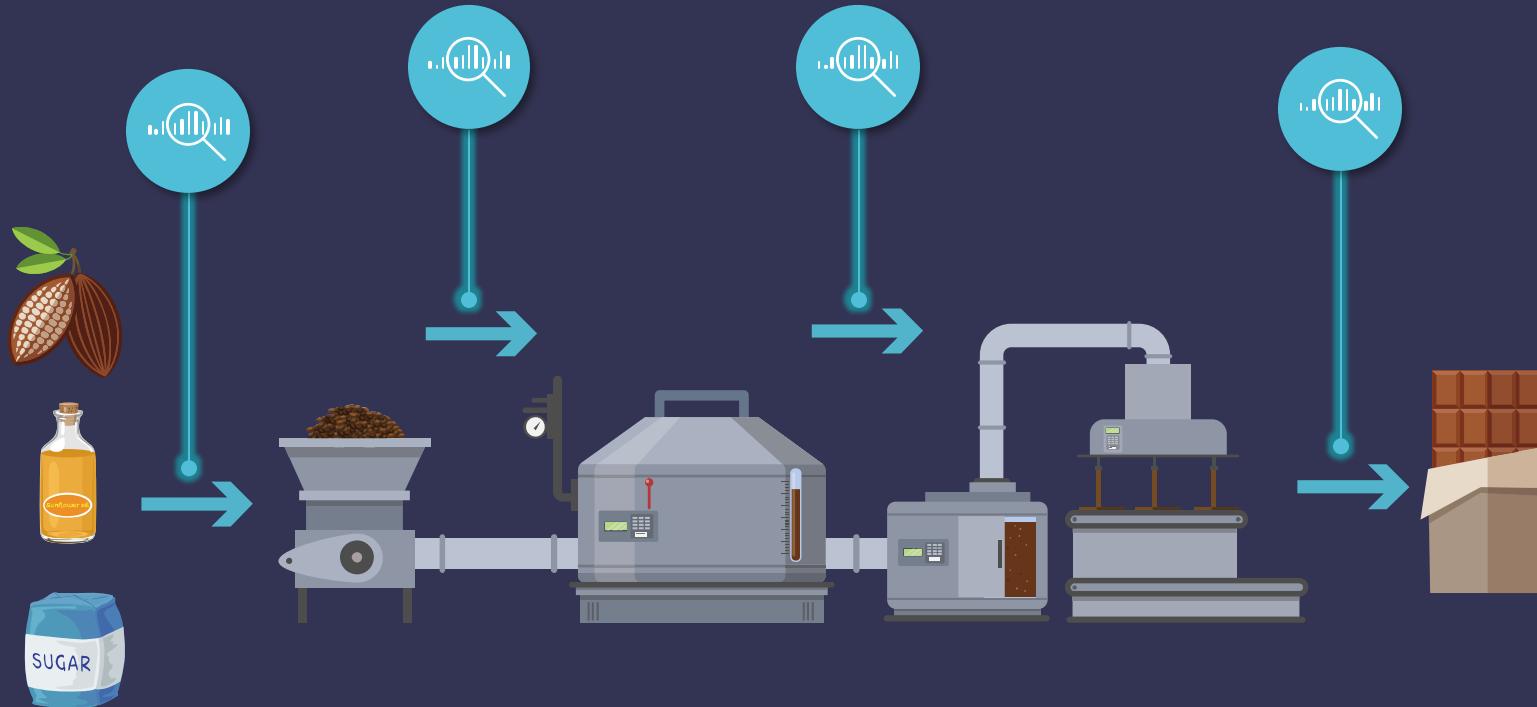
Lessons Learned

- A deep understanding of the app functionality, required data to enable functionalities as well as the use case is essential
- Some applications are very use case driven and base pretty much on already conducted loss-analysis





Mapping out plant line designs and clearly defining the data requirements are crucial steps to ensure data readiness



Challenges

- Smart decision on target scope for data points (trade-off between accuracy and efficiency)
- Different plants might have different application concepts, data points cannot be entirely generalized



Approach

- Set up data scope and identify strategic measuring of line performance
- Draw the complete plant line design → Prepare data model for line design → Perform the data requirement gap-analysis



Lessons Learned

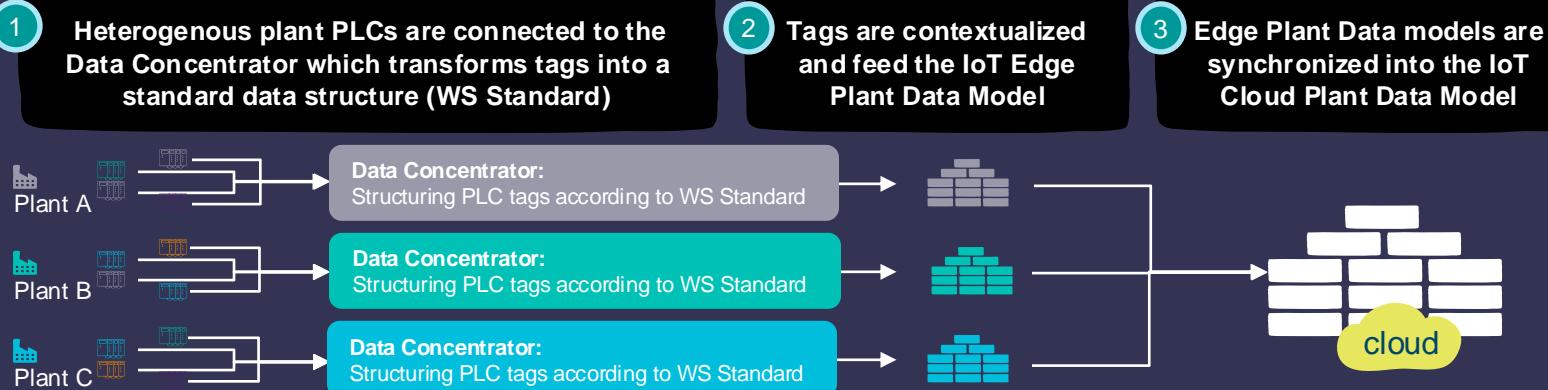
- A good, well-defined measurement ensure the quality and efficiency especially for long-term projects
- Requirement for standard objects so we do not build from scratch





Creating data templates to effectively track and manage diverse data streams ensures organized and accessible information management

OT Data (Operational technology) Data



Static Data

Static data refers to not frequently changing data. It provides further context to the user about plant-specific settings and for the automation data.

Plant Asset Hierarchy

Materials & Material Work

Unit Parameters

Reason codes

Shift Patterns

KPI Targets

Operations & Routings

Challenges

- Hardware purchases can be time-consuming and should be factored into overall planning
- Fundamental quality control is needed to avoid low level error and crash

Approach

- Before setting up the technical infrastructure and roll-out, it is necessary to assess aspects such as plant resources capabilities for OT data
- Install sensor → connect PLC → Computer level → aggregated data concentrator → cloud infrastructure → database Azure

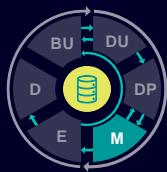


Lessons Learned

- Always stay focused on the goal of simplifying processes
- Ensure a backup resource for standard architecture setup and troubleshooting
- Onboarding the right data into the platform is effort intense and requires process knowledge and respective resource capacity



Developing four applications to empower production managers to make data-driven decisions



IOT platform

Challenges

- Standardization vs. individualization
- Integrating all aspects into four applications requires strong team collaboration



Approach

- All apps require additional configuration steps. These configuration data cannot be uploaded to the platform and must be set manually in the Configuration Management Portal.

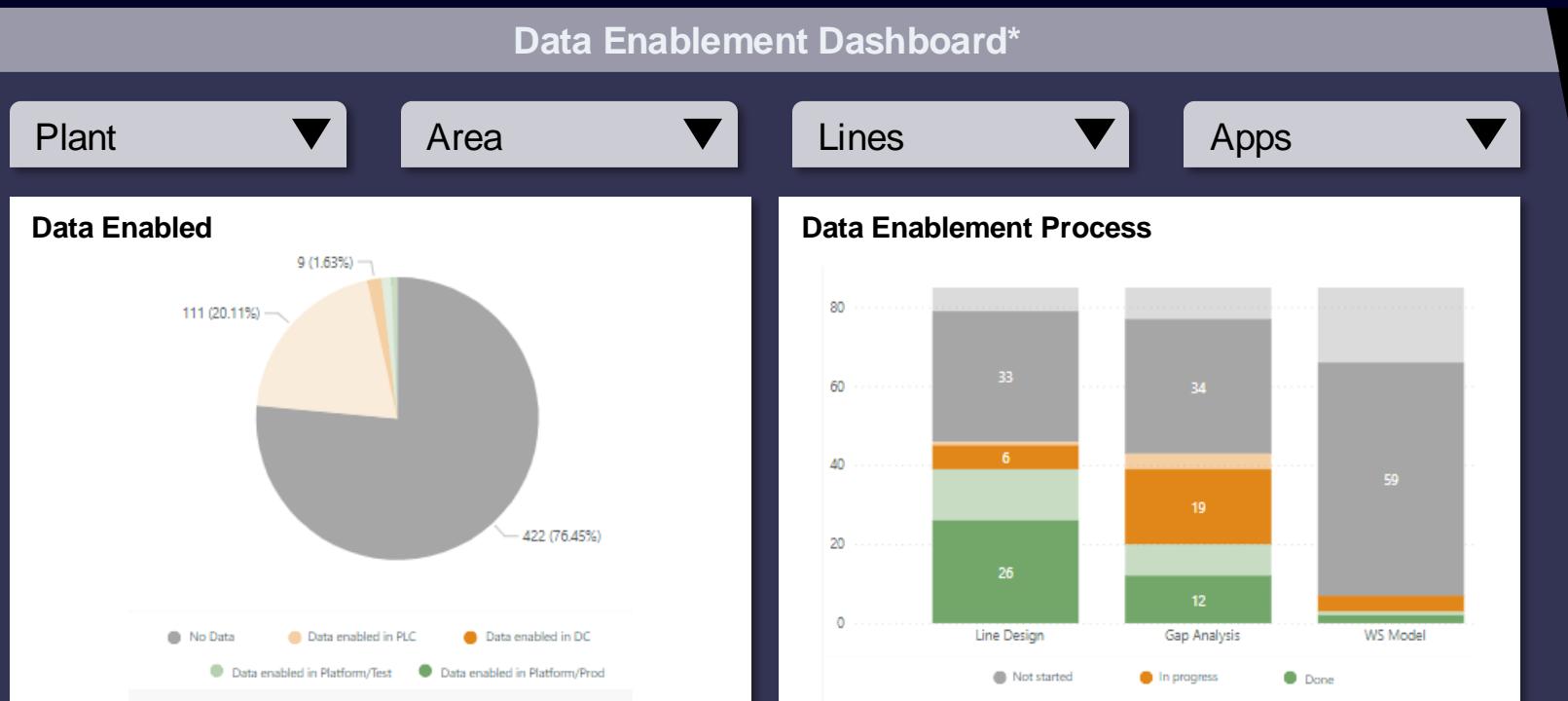


Lessons Learned

- Create solutions as individualized as necessary, but as standardized as possible! Adjustments for individual plants only if it results in high impact
- Stakeholder or end user should be involved as early as possible, they should be treated as a technical user for long term knowledge transfer



Designing dashboards to enhance visibility and understanding of data, ensuring full application functionality and transparency



* Simplified extracts of parts of the dashboard using dummy data

Approach

- Develop and deploy dashboards specifically designed to provide clear insights into application setup processes, enabling users to monitor configuration statuses, identify inconsistencies, and ensure all setup steps are completed efficiently and accurately



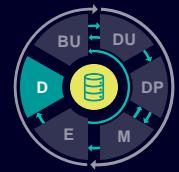
Lessons Learned

- Do what make sense → prioritize and manage effort
- Setting goals that are on the process level and don't need to go into details

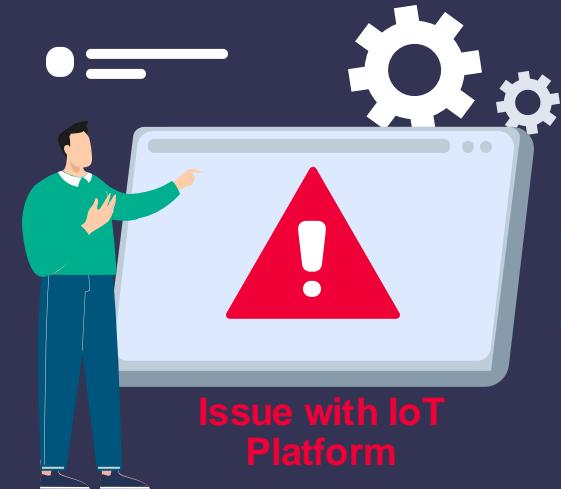
Challenges

- Large OT data sets making the process cumbersome and unable to track automatically (as they are not stored in the cloud, but gathered on local PLCs or edge devices)





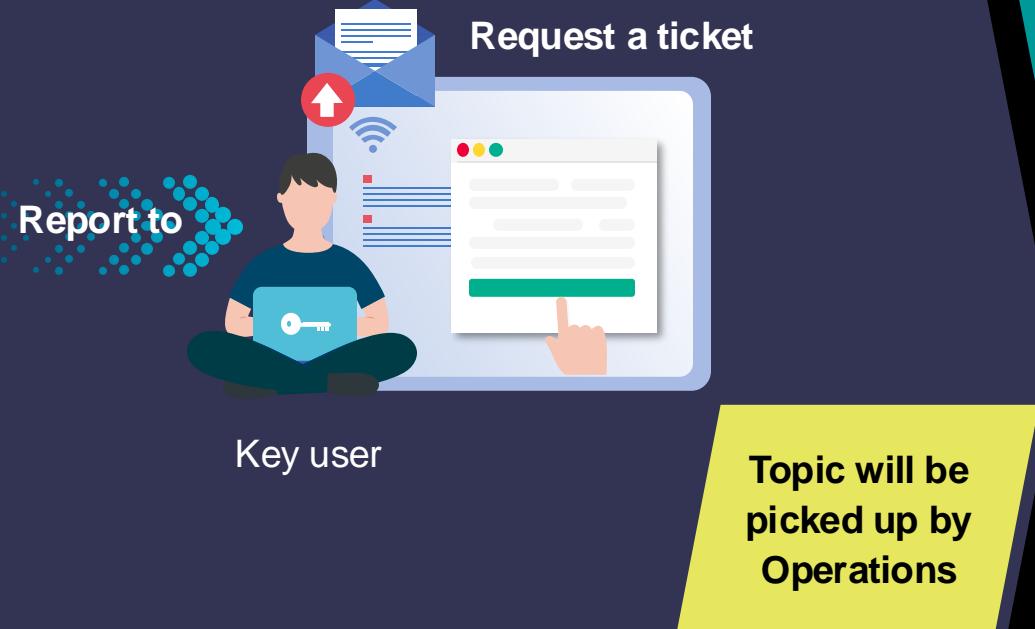
Performing in-system validation and troubleshooting to guarantee system stability and user satisfaction



End users detect

Challenges

- Integration into existing system
- User centric reporting cockpits



Approach

- Three stages (1) Go live preparation: Introduce the plant to the DevOps platform and the ticketing system, enable users to raise incidents in the platform (2) Go live week: Perform end-user trainings on-site at plant (3) Post go-live: The plant gets temporary support from M&T before transition to Operations

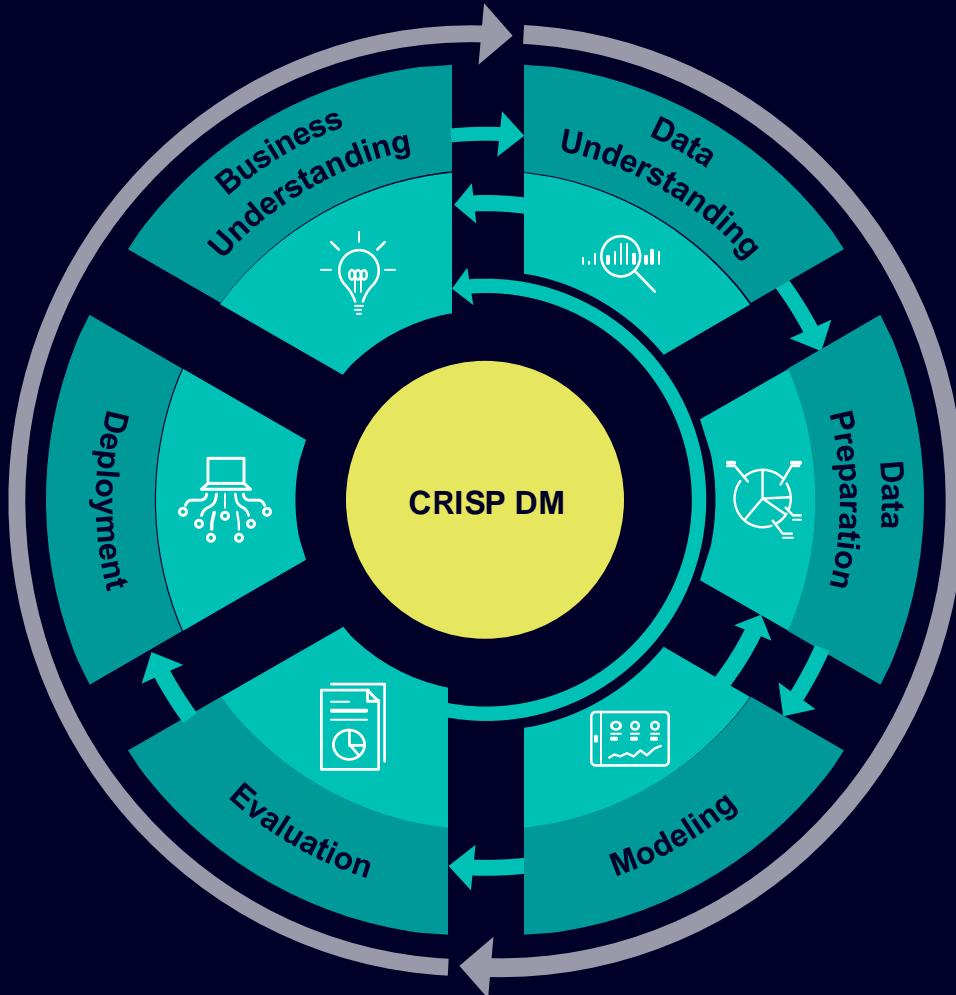


Lessons Learned

- Perform first Technical Go-Live in test environment and then move towards prod environment
- Transition from M&T support to Operations much more difficult as expected. It is important to have a clear cut and hand over responsibility and app ownership to plants.



The reasons for Data Science projects to fail are manifold – the Cross Industry Standard Process for Data Mining (CRISP-DM) is one lever to mitigate the risk of failure

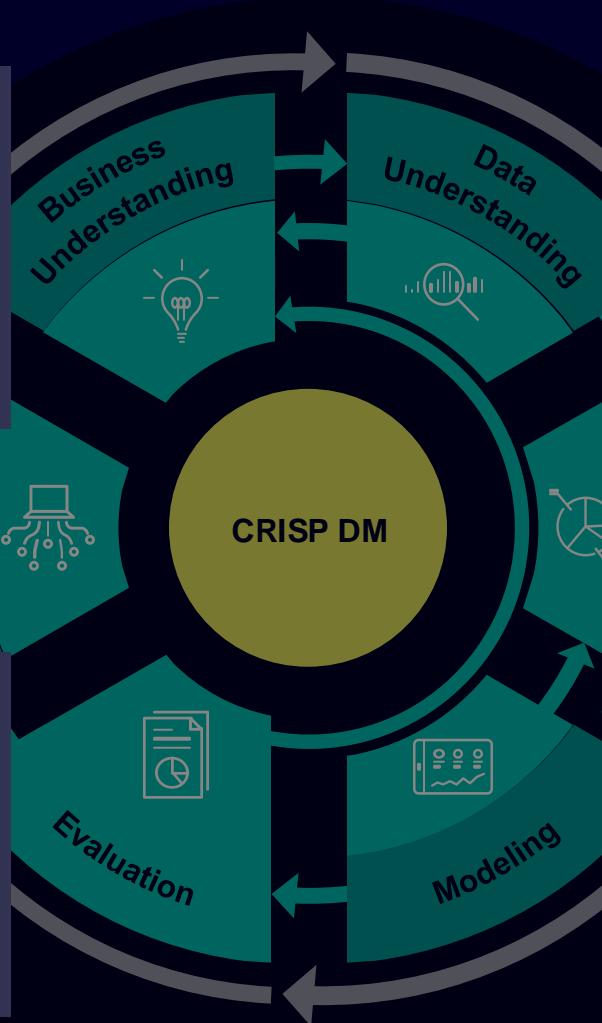


Source: CRISP-DM 1.0 (Chapman et al.)

The reasons for Data Science projects to fail are manifold – the Cross Industry Standard Process for Data Mining (CRISP-DM) is one lever to mitigate the risk of failure



Put yourself in your **customer's perspective** and understand the problem before developing an AI solution.



It is important to **monitor how models work with real data**. Fixes need to be made or the model needs to be updated.

While planning is important, most business projects do not go as originally planned. It is important to be **flexible** and **adjust the AI strategy** if necessary



Building an AI model is usually an **iterative process**. Systematic **documentation** of each experiment and the corresponding results helps to build organizational knowledge.



Questions



SIEMENS ADVANTA CONSULTING

Case Study Challenge

NOVA | March 11th, 2025

SIEMENS

The challenge is a monthly sales forecast based on real data from a Siemens business unit in Germany

Challenge overview

Sales forecast on monthly basis

Business area

- Selected product groups of one Business Unit of our Smart Infrastructure Division
- Focus on biggest country of business unit (Germany)

Data

- Sales data from October 2018 to April 2022
- Important macro-economical indices

Evaluation:

- Quantitative evaluation of score (Metric: RMSE) via separated test set (May 2022 to February 2023 – 10 months)
- Submission until March 31st via Moodle and email

Why is AI-driven sales forecast so important?



Manual forecasting is highly resource intensive – many “person-days” per month



Manual forecasting is biased through aggregated “judgement” of multiple stakeholders



Information scattered over many data sources ever-changing base-lines



“Opportunity Cost” of poor forecasting is significant – on working capital &/or customer satisfaction

Businesses and Services of Siemens AG

Industrial businesses

Digital Industries



Smart Infrastructure



Mobility

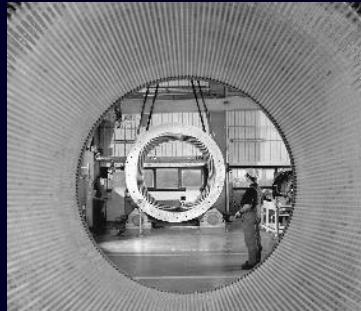


Siemens Healthineers¹



Other businesses

Portfolio Companies



Siemens Advanta



Siemens Financial Services



Services

Siemens Real Estate



Global Business Services



¹ Publicly listed subsidiary of Siemens; Siemens' share in Siemens Healthineers: 75%

Siemens SI EA Germany Product Portfolio for Switch Gear Products in Smart Power Distribution from Medium-voltage to Low-voltage

Medium-voltage Components, Systems & Solutions

Components



Primary Air-insulated Switchgear



Gas-insulated Switchgear Prim / Sec



SF6-free blue GIS Prim / Sec



Generator Circuit Breaker Switchgear



Photovoltaic (PV inverter systems)



Low-voltage Systems

Power Distribution Boards & Motor Control Centers



Automation, Protection & Communication for high- and medium-voltage Systems

Protection



Automation



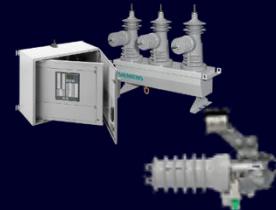
Power Quality



Smart Communication



Outdoor Distribution Systems



Solution Business (E-House systems)



Busbar trunking systems



Services

Data provided consists of daily sales data from Siemens product groups and key market indices



Sales data

- Daily sales data per GCK (product groups) in EUR
 - Training set (daily): 'Sales data.csv'
 - Test set (monthly): 'Test Set Template.csv'



Macro economical data

- Important macro-economic indices for Siemens in its most important countries: 'Market data.xlsx'
 - This includes for example:
 - Production Index Machinery & Electricals
 - Shipments Index Machinery & Electricals
 - Price of Base Metals
 - Price of Energy
 - Price of Metals & Minerals
 - Price of Natural gas index
 - Price of Crude oil, average
 - Price of Copper
 - Producer Prices

Datasets available on Moodle

Dataset available on Moodle

The test set should be sent by March 31st and the results will be determined quantitatively by SAC



Submission of results

- **Deadline**
 - Submission until 23h59 March 31st , 2025
- **Test set submitted also to Siemens Advanta Consulting**
 - **Mail addresses:** julian.spall@siemens.com
 - **Subject:** NOVA/SAC – Group Name (e.g., NOVA/SAC, Group A)
 - **Format:** Please use the CSV 'Test Set Template' including three columns (Year Month, Mapped_GCK, Sales EUR)



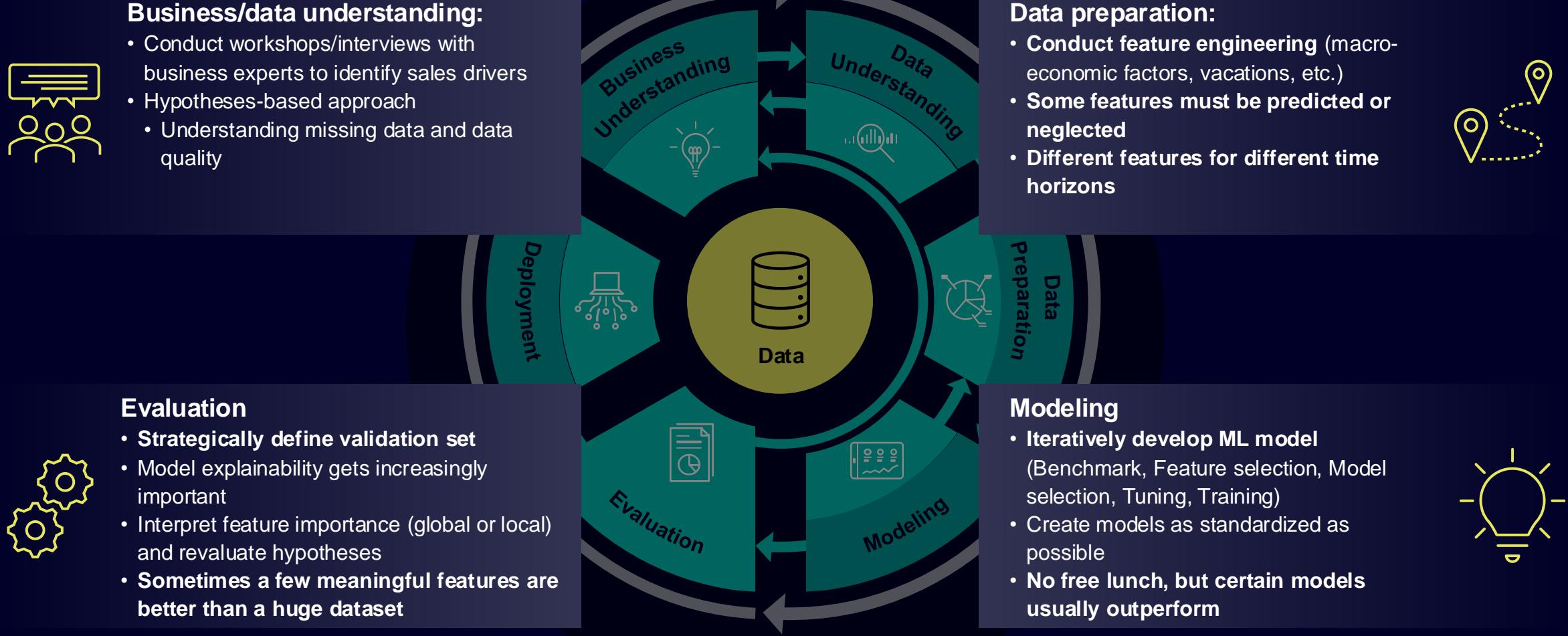
Evaluation of results

- **Quantitative evaluation of results by SAC**
 - Score on test set via RMSE
- **Qualitative evaluation of results, presentation and code by NOVA IMS**

Questions



The reasons for Data Science projects to fail are manifold – the CRISP-DM framework is one lever to mitigate the risk of failure



Source: Siemens Advanta Consulting

Job Opportunities with Siemens in Lisbon Portugal



Siemens Mobility

Siemens Mobility is an independently managed company of Siemens AG and has been a leading provider in the field of mobility for over 160 years. Our core business includes rail vehicles, rail automation and electrification solutions, turnkey systems, and associated services. Through digitalization, we make infrastructures intelligent and create opportunities that sustainably and seamlessly take us from A to B. Our 39,800 employees are pioneers in mobility, helping to keep the world moving.

Data Analyst (m/f/d)



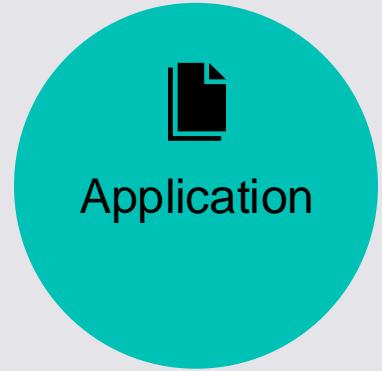
Senior Data Scientist (m/f/d)



We offer a three-stage interview process with timely feedback loops

Overview of our interview process

01



Application

Response within 14 days

02



Data Analytics (Senior) Consultants

1st round with Data Analytics Team

1 Interview (Mix of Behavioral Questions & Analytical part/case)

2nd round – Interview day

3 Interviews (Mix of Behavioral Questions & Case interviews)

Siemens Advanta Consulting will provide you with personal feedback

03



Contract Offer

Post Offer Process

Connect with us on LinkedIn!



Julia Spall
Senior Consultant



Justus Hümmer
Senior Consultant



Ida Manko
Recruiting Manager



Laura Ponschab
Recruiting Manager

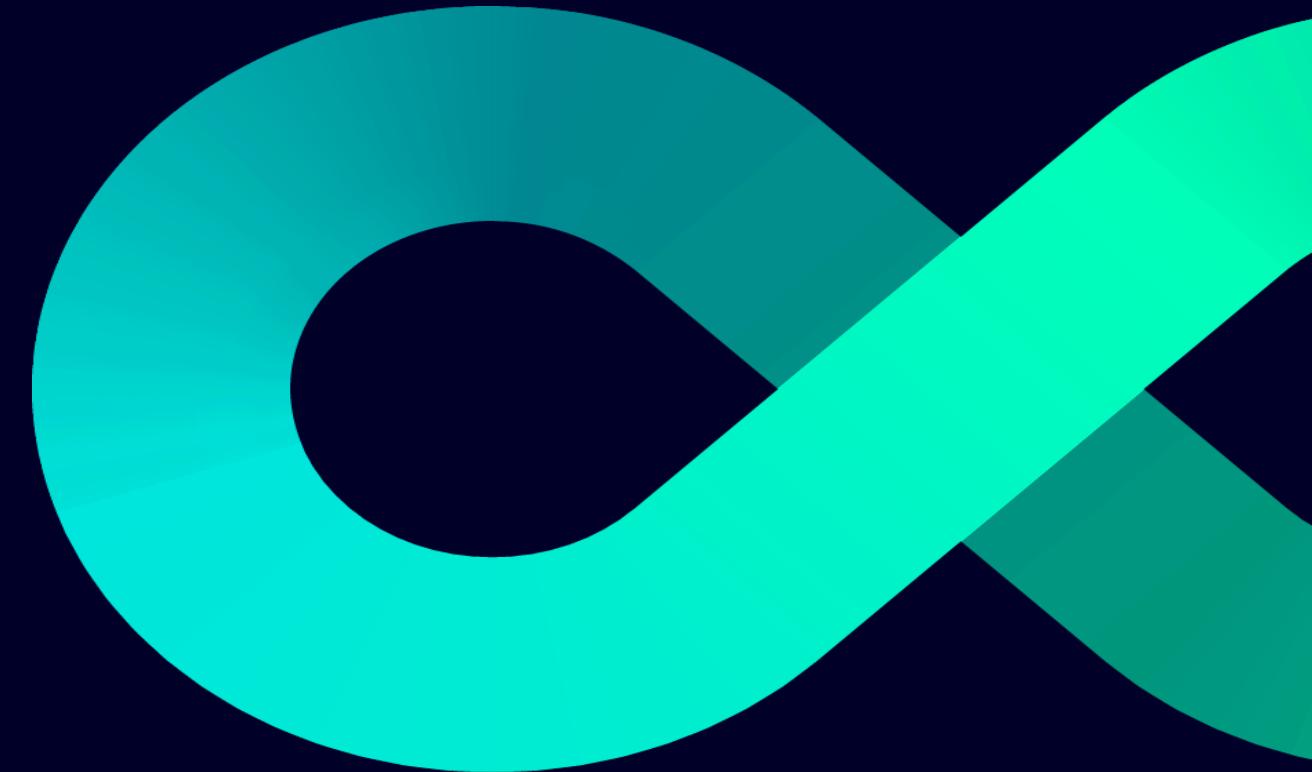


SIEMENS ADVANTA CONSULTING

Virtual Goodie Bag



By scanning the QR code, our Virtual Goodie Bag awaits you. Inside, you'll find exciting content. Enjoy exploring!



Thank You

& have a great day!

