

SFS Americas Digitalization Learning Program: Artificial Intelligence @ work @ SFS

SFS IT&CYS AI

Dr. Ulrike Dowie, Markus Dressel, Dr. Konstantinos Trochidis, Rahatu Alasane,
Detlef Schuetz & Diogo Alves

AI Chapter – who are we?



**Dr. Ulrike
Dowie**

Team Lead & Facilitator
FDSN Co-Founder



**Markus
Dreßel**

Senior IT Solution
Expert for AI



**Detlef
Schuetz**

Senior Solution
Architect for AI



Konstantinos Trochidis
Lead ML Engineer



**Rahatu
Alasane**
Data Scientist and
ML Engineer



**Henrique
Silva**

Machine Learning
Engineer Trainee



**José Filipe
Brito de Jesus**

Machine Learning
Engineer Trainee



**Diogo Goncalo
Dias Alves**

Data Scientist and
ML Engineer



**Doris Etonam
Alifui-Segbaya**

Communications



**Marika
Horne**

Communications

AI team: Business value & User experience go first!

Our mission: We enable you for data-driven decision making by applying Machine Learning expertise and the latest AI technologies and **jointly improve processes & enable data-driven financial services.**



See some User statements on AI solution „DALE“:

”

We often class certain types of documents we examine as 'high volume but low risk'. Our business partners will ask us at Legal for help when they are uncertain about something or when they want to have another pair of eyes review the documents.



Tom Lau
Senior Legal Counsel,
Siemens Bank

This review drags out the process and rarely turns up any truly critical risks. This is why we started to search for a way to help our colleagues conduct an automated review of the documents.

”

DALE ensures no interruption to the current application landscape for its users, as it is build as a Microsoft Office Plug-in. No additional installation is required so long as access is provided. In the back-end, the application is hosted onto Platform-as-a-Service Eagle Openshift, which allows us to quickly develop, host and scale our application.



Jolijn de Kock
SFS IT Project
Manager

”

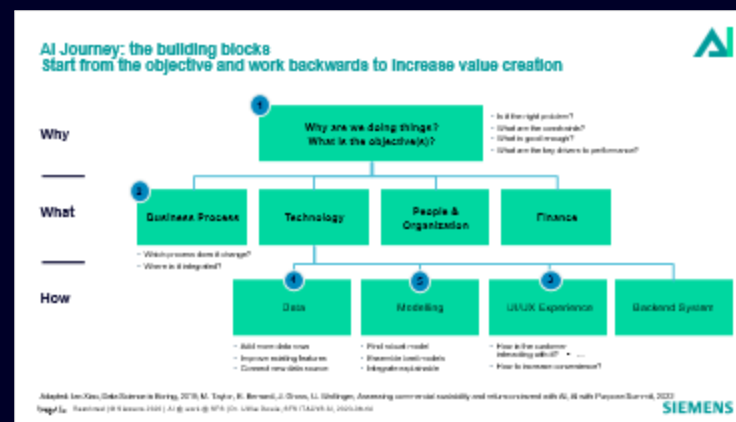
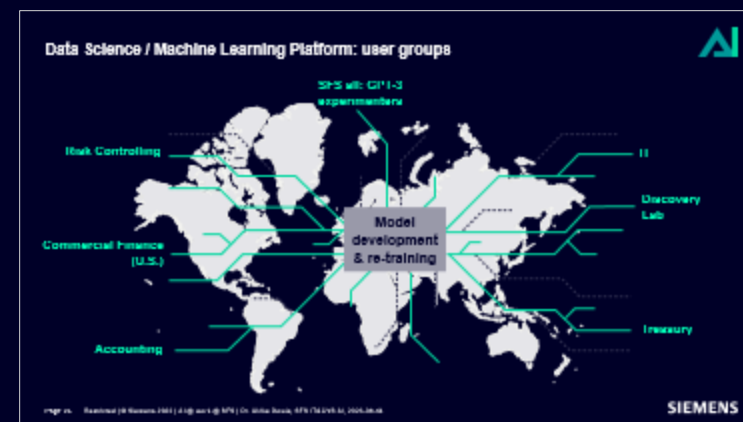
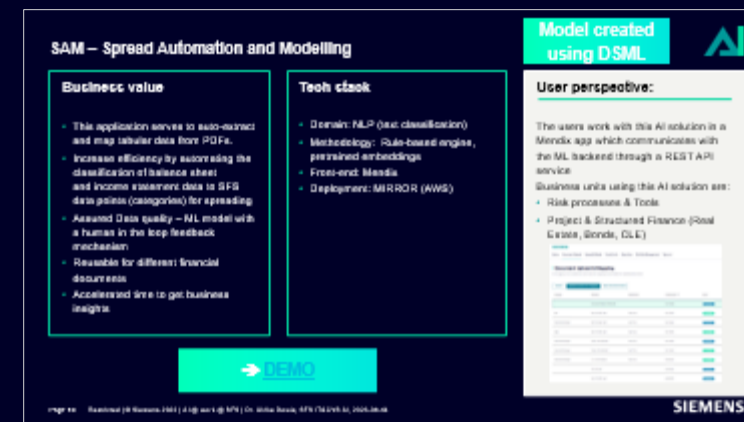


Martina Adamek
Senior Riskmanager,
Siemens Bank

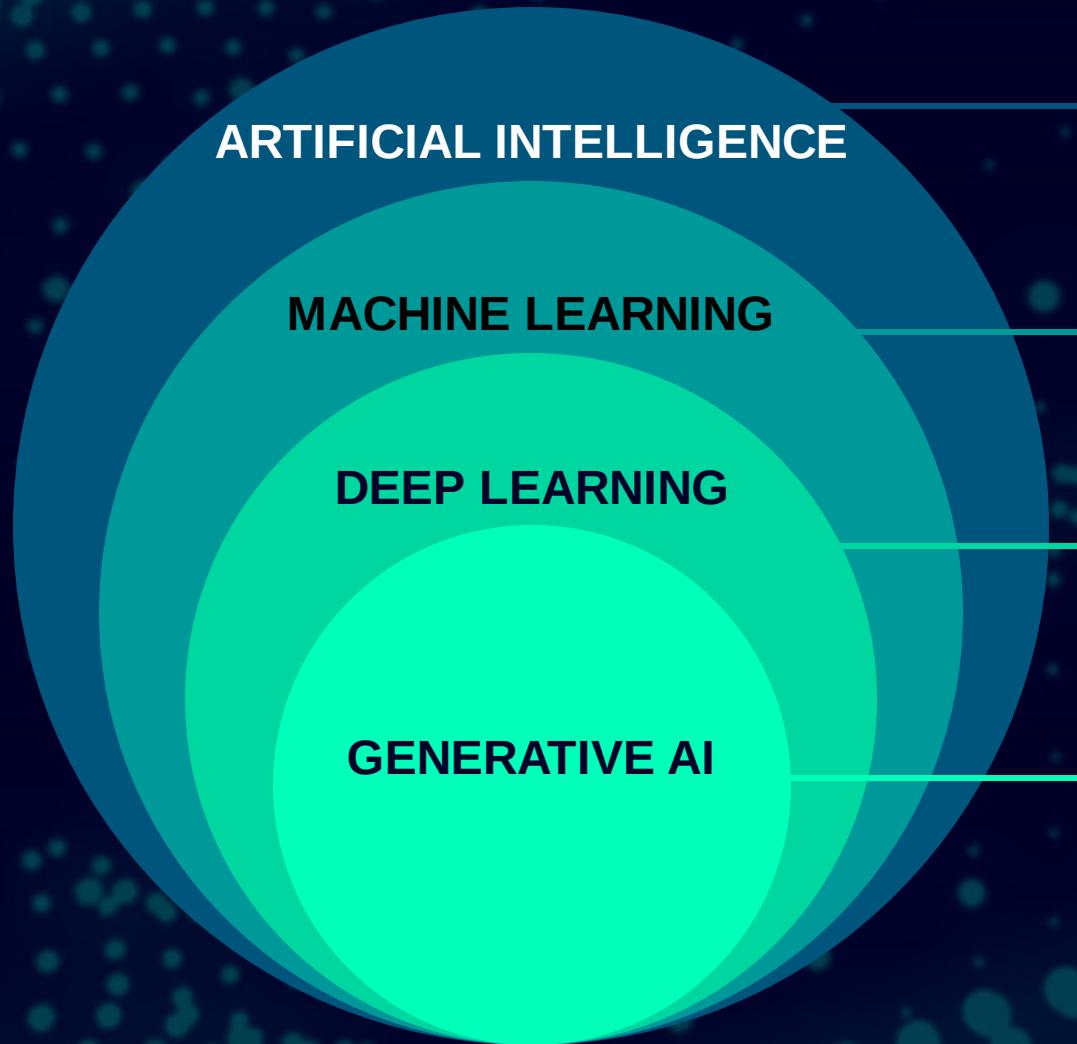
I think DALE is really very intuitive and easy to use. You simply let it work on the side, and you basically have a second pair of eyes reviewing the document. The tool gives me the assurance that I need. I no longer have to ask a colleague from Legal every time I have a legal question.

Table of contents

Summary zoom



What is AI?



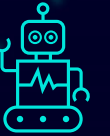
Artificial Intelligence

The field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence



Machine Learning

Subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions



Deep Learning

A machine learning technique in which layers of neural networks are used to process data and make decisions



Generative AI

Generative AI algorithms (such as ChatGPT) can be used to create new written, visual and auditory content given prompts or existing data.



***~80% of current AI research is focused on Generative AI.**

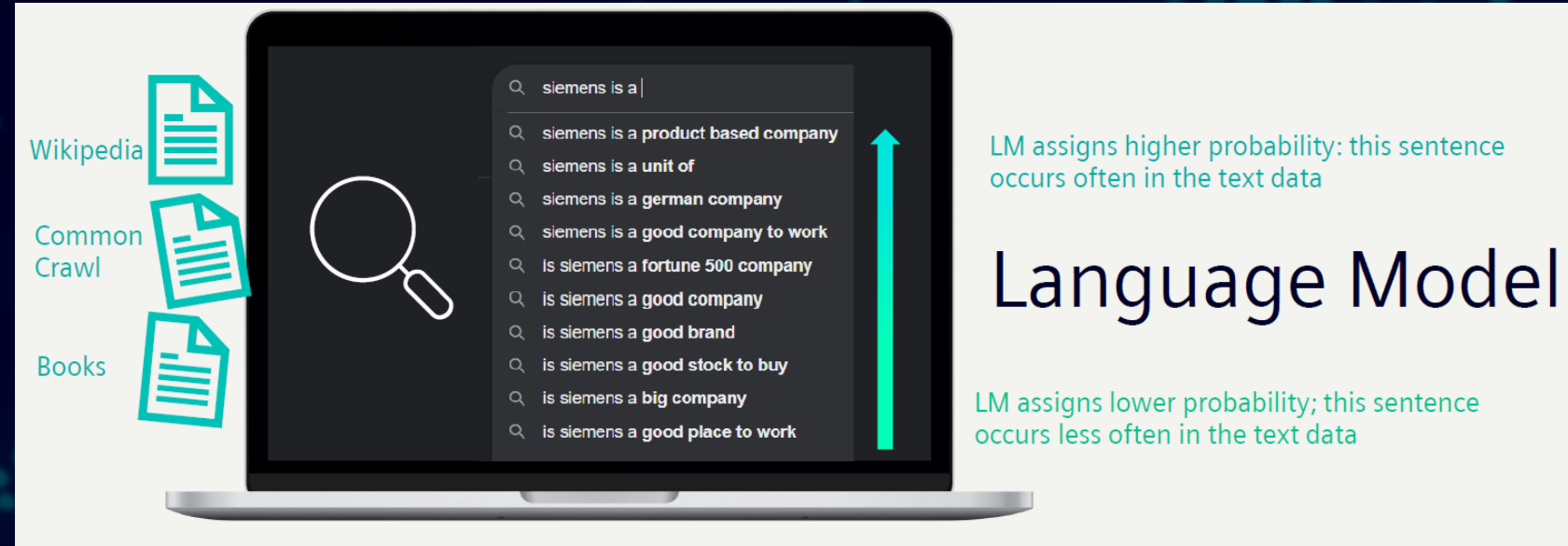
*<https://www.economist.com/interactive/briefing/2022/06/11/huge-foundation-models-are-turbo-charging-ai-progress>

ChatGPT is a Large Language Model – What’s the magic behind it?

Scientific Explanation



Probability distribution over word sequences. Given any sequence of words of length m , a language model assigns a probability $P(w_1, \dots, w_m)$ to the whole sequence.



The **Language Model** outputs the most probable next word(s) and **generates** in this manner long text. Its input, “siemens is a” is called **prompt**.

Writing prompts to fulfill a certain task, such as rewriting an e-mail in a polite way, is called **prompt engineering**.

What are the top applications of LLMs like GPT?



Text Classification

Classify textual content into user-defined categories

SFS Example:

Identify risks in annual reports



Chatbots

Intelligent question and answer system that is able to understand questions and simulate human conversation

SFS Example:

QA chatbot on financial & tax documents



Content Creation

Generate human-like text, making them useful for creating content such as emails, marketing material, presentations and even images

SFS Example:

Writing support for Comms & Customer Service



Information Extraction

Extract entities and their relationship from documents to accelerate business processes.

SFS Example:

Knowledge extraction from financial documents



Code Assistant

Write new code, find bugs, translate from one programming language to another

SFS Example:

Software Dev CoPilot for standard programming languages



Sentiment Analysis

Extract the sentiment to proactively steer your customer services team, start marketing initiatives or identify new opportunities for your business

SFS Example:

News Sentiment Analysis

Ruled-based approach vs. Machine Learning approach



vs.



**Rule-based
approach**



Rules-generated by human
Empirically derived rules
NO labeled data needed
Low flexibility to generalize

**Machine Learning
approach**



Feature extraction by machine to detect patterns
Automatic generated rules by the model
MUCH labeled data are needed
High flexibility to generalize

What is AI capable of and what not



SIEMENS

Workspaces Documents Settings

–

Reset

+

⏮

⏪

⏩

⏭

137 of 358

⏴

⏵

⏶

⏷

⏸

Notes unconfirmed

Document status: ✓
Notes sync status: ✓

20.1 Financial statements

The Borrower shall supply to the Intercreditor Agent:

- 20.1.1 as soon as the same become available, but in any event within ten (10) Months after the end of each financial year of the relevant Obligor, the audited annual financial statements of each Obligor for that financial year;
- 20.1.2 as soon as the same become available, but in any event within one hundred and twenty (120) days after the end of each half of each financial year of the relevant Obligor, the unaudited financial statements of each Obligor for that financial half year; and

Detect a specific clause in a document based on 1000 examples

mark'd Search Link Collection Logout

03. CTA - Neart Na Gaoithe Offshore Wind - 19 Sep 2019 (PRJ).pdf (DALE2 - Phase 1)

Notes

Y Hide filters

Group by category Collapse all Expand all

Categories

Search text here

73 of 73 total

Undertaking: Financial Information

8 of 8

20.1.1 as soon as the same become available, but in any event within ten (10) Months after the end of each financial year of the relevant Obligor, the audited annual financial statements of each Obligor for that financial year;

100.00%

20.1.2 as soon as the same become available, but in any event within one hundred and twenty (120) days after the end of each half of each financial year of the relevant Obligor, the unaudited financial statements of each Obligor for that financial half year; and

100.00%

The Borrower shall supply a Compliance Certificate to the Intercreditor Agent on or promptly after each Scheduled Repayment Date substantially in the form set out in Schedule 9 (Form of Compliance Certificate).

100.00%

The Borrower shall no later than sixty (60) days prior to the forecast Commercial Operations Date deliver to the Intercreditor Agent, as electronic copies for the Senior Lenders and the Senior Lenders' Technical Adviser, a draft operating budget, prepared by it with due care for the first financial year following the forecast Commercial Operations Date.

100.00%

21.3.1 During the Construction Period, the Borrower shall, no later than thirty (30) days (or such later date agreed with the Intercreditor Agent) after the end of each calendar month (commencing from the calendar month following Financial Close until including the calendar month in which the Final Completion Date occurs) deliver to the Intercreditor Agent and the Senior Lenders' Technical Adviser, as electronic copies for the Senior Lenders and EKF, a report prepared by it with Good Industry Practice (the "Construction Report") setting out:

100.00%

During the Operating Period, the Borrower shall, within thirty (30) days after each Scheduled Repayment Date, deliver to the Intercreditor Agent as electronic copies for the Senior Lenders and EKF:

100.00%

21.5.1 an operating report prepared by it with due care, for the half-year ending on that Scheduled Repayment Date including details as to:

100.00%

The Borrower shall prepare in consultation with the Intercreditor Agent, and deliver to the Intercreditor Agent, as electronic copies for the Senior Lenders,

100.00%

Detect a specific clause in a document based on 10 examples

Where AI works and where it does not?



AI works well when..
A lot of data are cleaned and available
There are distinct patterns within the data

AI does not work well when..
The data are incorrect and limited
Environment is changing radically
There are no distinct patterns within the data

AI areas of application



1 Natural Language Processing

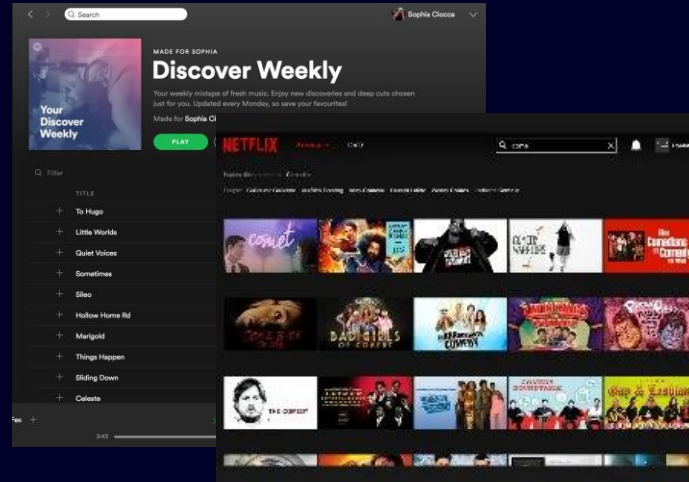
Algorithms that can support automatic **categorization** of documents based on type and **extraction** of specific information from the documents such as name, origin, company, revenue.



NLP

2 Recommender Systems

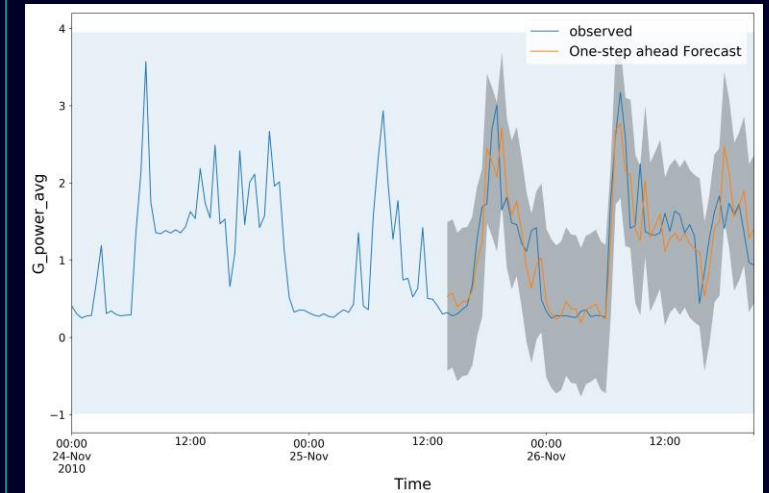
Suggesting new content such as music, product or articles to a user based on specific preferences and previous behavior.



Recommendation

3 Time-series predictions

Create and fine-tune models to predict future values based on previously observed values.



Time-series analysis

AI areas of application: what's cooking?



Tasks today:

1. Outlier detection
2. Customer mapping: Are these 2 customers the same? Or belong to the same corporate customer?
3. Forecasting of NCO, other KPIs
4. Data Cleansing (of asset portfolio data)
5. Ad hoc analyses on tabular data
6. Extract trends and other info from Financial Outlook Reports
7. Guideline Search
8. Insurance Certificates: extract unstructured information from lengthy documents
9. Automate the (tedious) booking process



Value Proposition:

1. Focused search, reduced manual work
2. Data quality increased, reduces manual work
3. Get unbiased Forecast values to support your planning, reduce time
4. Reduced manual work, improved data quality
5. Faster reporting answers, reduced manual work
6. – 9. reduce manual work and time needed, make processes more efficient

AI areas of application: Newsroom (experiment done by Discovery Lab)

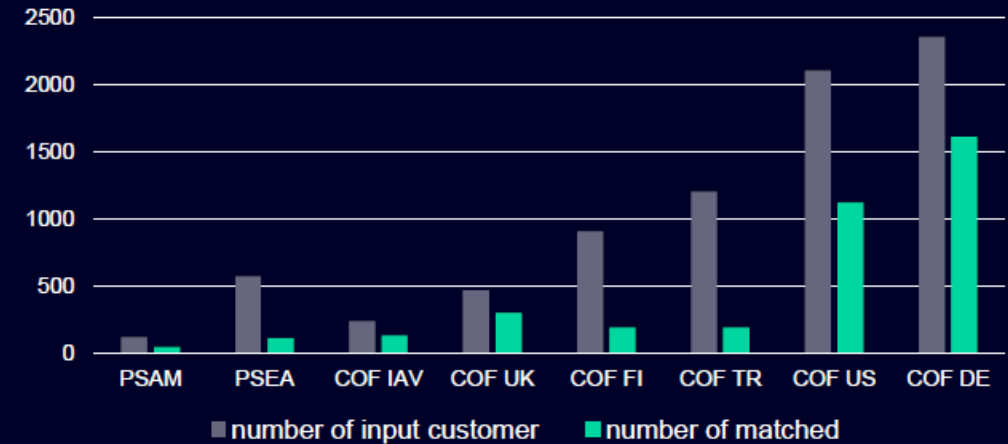


Can we find news regarding our customers and map them accordingly

- With certain search criteria and matching algorithms, we have been able to match a significant amount of our customers to news articles

Can we filter news articles based on the topics they deal with to identify relevant articles for our business

- We created a toolbox to calculate sentiment and relevancy scores for identified news article in order to filter relevant news (identify “needle in the haystack”)



Personas:

1 COF Direct Sales



Value Proposition:

- provide triggers to start new sales activities early
 - provide triggers to start sales activities at existing customers
 - provide summarized reports of relevant trade press for keeping up-to-date about industry trends
-
- Clustered news overview related to vendors, competitors, industry developments
 - provide summarized and clustered information / reports for keeping up-to-date about industry trends



AI areas of application along the Financing Process

1 Trend identification & simulation

Mine vast amount of unstructured data from economic reports, company earnings, news, social media, and other trends that can impact the investment portfolio, or simulate conditions to assess a potential impact on it.

Classification

4 “Just in time” lending

Apply AI techniques to accelerate the risk assessment process for credit applications

Classification

6 Compliance assurance

Analyze structured transactions in conjunction with unstructured data such as documents, emails, or voice orders to identify internal practices comprising compliance.

Classification

2 Dynamic Pricing

Apply machine learning to broader data assets to provide more accurate pricing and optimize underwriting. E.g., Pay-per-Use based on IoT, or price recommendations based on historical and market data

Optimization

(Pre-)Sales &
Origination

1 2 3



Risk

4 5



Operations

6 7 8

7 Analytics-enabled collections

Use power of advanced analytics and machine learning to transform collections to accelerate development of treatment and contact strategies.

Recommendation

3 Personalized Financial Services

Provide customized services such as portfolio management or tailored incentives using AI-powered recommendation techniques, often also accessed through conversational AI.

Recommendation

5 Default prediction

Improve accuracy of risk assessment, which can be used to increase revenue by issuing loans that previously would have been denied or by anticipating and avoiding defaults.

Classification

8 Fraud Detection

Use AI to spot unusual patterns by analyzing multiple data points to identify fraudulent transaction that rule-based analysis may miss.

Classification

SAM – Spread Automation and Modelling

Business value

- This application serves to auto-extract and map tabular data from PDFs.
- Increase efficiency by automating the classification of balance sheet and income statement data to SFS data points (categories) for spreading
- Assured Data quality – ML model with a human in the loop feedback mechanism
- Reusable for different financial documents
- Accelerated time to get business insights

Tech stack

- Domain: NLP (text classification)
- Methodology: Rule-based engine, pretrained embeddings
- Front-end: Mendix
- Deployment: MIRROR (AWS)

→ [DEMO](#)

Model created
using DSML



User perspective:

The users work with this AI solution in a Mendix app which communicates with the ML backend through a REST API service

Business units using this AI solution are:

- Risk processes & Tools
- Project & Structured Finance (Real Estate, Bonds, CLE)

The screenshot shows the Siemens 'Document Upload & Mapping' interface. It includes a navigation bar with links like Home, Document Upload, Spread & Model, Trend Cards, Reporting, Portfolio Management, and Sign out. Below the navigation bar, there's a section titled 'Document Upload & Mapping' with a sub-header 'In this page you can upload new documents for mapping and see details for uploaded documents.' There are three buttons: 'Search', 'Upload document for mapping', and 'Open document details'. Below these buttons is a table with columns: Company, Filename, Sales Branch, Created Date, and Status. The table contains several rows of data, including 'Example of Balance Sheet.pdf', 'Del 10K 2021.pdf', 'Advanced Drainage', 'Mita', 'Heraeus 10K 2020.pdf', 'Yum 10K 2020.pdf', 'ADS 10K.pdf', and 'Del 10K 2021.pdf'. Each row has a 'Status' column with a button labeled 'Completed'.

Company	Filename	Sales Branch	Created Date	Status
	Example of Balance Sheet.pdf		6/21/2022	Completed
Del	Del 10K 2021.pdf	Cash Flow	6/21/2022	Completed
Advanced Drainage	Del 10K 2021.pdf	Cash Flow	6/21/2022	Completed
Mita	Del 10K 2021.pdf	Cash Flow	6/21/2022	Completed
Advanced Drainage	Heraeus 10K 2020.pdf	Cash Flow	6/21/2022	Completed
Advanced Drainage	Heraeus 10K 2020.pdf	Cash Flow	6/21/2022	Completed
Advanced Drainage	Yum 10K 2020.pdf	Cash Flow	6/20/2022	Completed
	ADS 10K.pdf		6/20/2022	Completed
	Del 10K 2021.pdf		6/20/2022	Completed

DALE2.0 – Digital Advanced Legal Emulator

Model created
using DSML



Business value:

- Increase efficiency by automating the identification of clauses as assistance for subject matter experts in credit facility agreements
- Keeping up with market standards in turnaround time for credit decisions
- Integration into current toolchain for document review (MS Word)
- Flexibility to support and scale document automation by transferring existing ML models to processes further document types

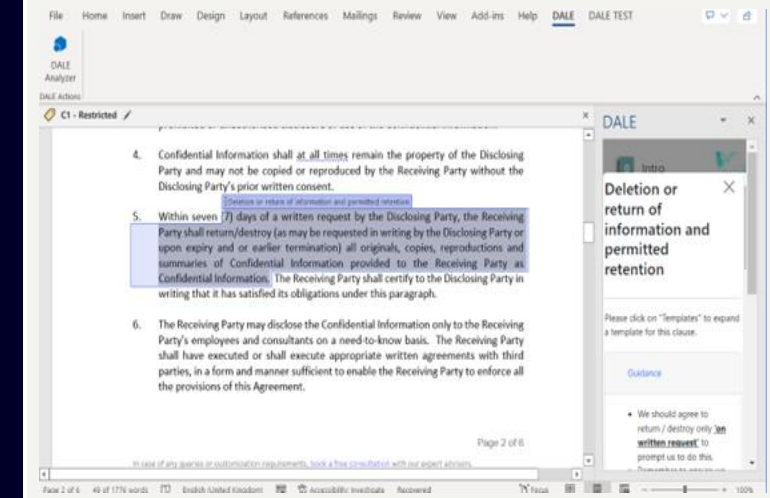
Tech stack:

- Domain: NLP (text classification)
- Methodology: Machine & Deep Learning (Random Forest, BiGRU, bert)
- Annotation Platform: Mark'd
- Front-end: React
- Deployment: OpenShift

User perspective:

The users access / work with this AI solution via a MS word plugin
Business units using the AI solution are:

- Risk processes & Tools
- Corporate Lending
- Banking and finance



[→ DEMO \(3 min\)](#)

[→ DEMO \(11 min\)](#)



MLDD - Machine Learning Due Diligence

Business value

- Accelerated due diligence process for the financing of energy projects
- Time savings in reading documents
- Standardization of risk classification
- Quality improvements in risk classification

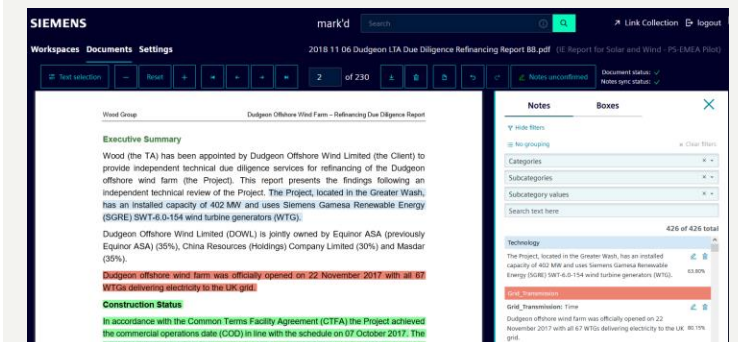
Tech stack

- Domain: NLP (text classification)
- Methodology: Machine Learning (Deep Learning), pretrained embeddings (FastText)
- Front-end: mark'd
- Deployment: AWS

User perspective:

MLDD is used in combination with mark'd as frontend

- User uploads documents into mark'd
- The prediction pipeline is automatically triggered by mark'd
- Text data are preprocessed, risk categories are assigned on sentence level and the result is pushed back to mark'd
- The user can review the risks in mark'd



→ DEMO (6 min)



LENDING APP

Business value:

- Seamless loan servicing: loan notices are auto-identified, invoices generated and sent, and incoming payments matched against Loan receivables
- Saves time and operational costs by leveraging AI and automating repetitive tasks.
- Handling of high loads of incoming notices possible without losing oversight.
- Lean end to end process within the Lending team enabled with the use of a single user interface.

Tech stack

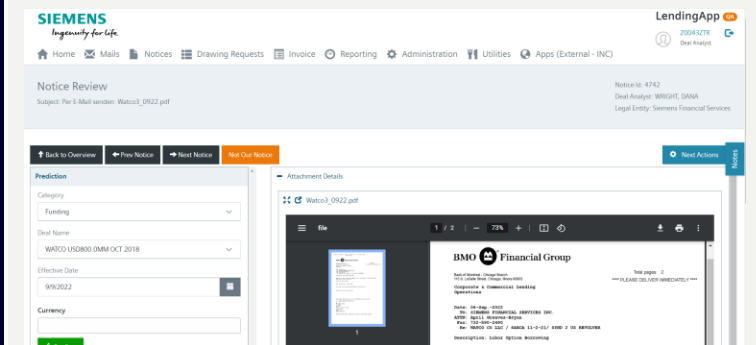
- Domain: NLP (Document classification, extraction of key figures)
- Methodology: Machine Learning, Neural network, Random Forest
- Front-end: Mendix
- Deployment: OpenShift

User perspective:

The prediction service is accessed via the Mendix frontend.

- Currently used by Risk Controlling and Operations in the United States.

 **DEMO**



PARDON

Business value

- An AI model calculates the probability of default used for scoring and auto-decisioning.
- Reduced number of credit hits
- Higher volume of small tickets can be processed with the same team size.
- Increased partner satisfaction due to lower cycle times.
- More precise pricing and risk capital calculation.

→ **DEMO (User Interface)**

Tech stack

- Domain: Default prediction
- Methodology: Machine Learning (XGBoost)
- Front-end: PowerCurve
- Deployment: MIRROR (AWS)

→ **DEMO (ML Backend)**

**Solution developed by
Risk experts / RCA**

User perspective:

The PARDON AI service is accessed via PowerCurve (Credit Decision tool).

- Currently used by Commercial Finance in UK and Nordics.
- Started implementation for new country (Spain)

The screenshot displays the PARDON AI service interface, which is used for automated rating and decision-making. It features a 'Rating' section with a table of 'Top 10 Rating Factors' and their corresponding scores. The 'Automated Rating' section shows a 'Scoring Rating' of 5- and a 'Calculated Probability Of Default' of 0.01033. The 'Decision' section includes a 'Manual Decision' dropdown set to 'Choose One', a 'Decision Valid to' date of 2021-11-25, and a 'Reason' field. A '10% Overdraft' checkbox is also present.

Top 10 Rating Factors	
ProportionCurtDirToPrevDirLast12	0.87858
CompanyOBT	-0.34193
ScoringShareholdersFunds	0.26815
LatestAccountsAge	-0.23304
WorkingCapital	-0.19893
Ratio_gross_assets_curr_liabil	-0.18540
CommDelphScore	0.18144
PreviousAddressDateChanged	-0.17682
GrossAssets	-0.13563
TimeRelationMonths	-0.12916



mark'd – Text annotation platform

Business value

- Documents for NLP use cases can be stored in central location
- Data are extracted and preprocessed in a standardized way
- It serves as front-end for NLP algorithms
- Results from NLP algorithms are visualized in the documents and can be reviewed by users
- Fully owned Siemens product
- Compliant with Siemens security requirements

Tech stack

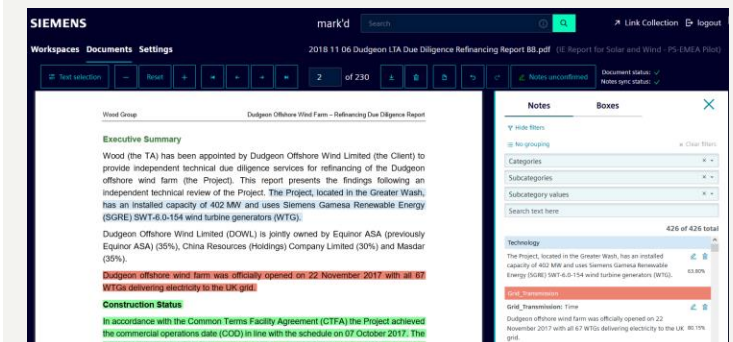
- Backend
 - Java
 - Spring Boot
 - Hibernate
- Front-end
 - Typescript
 - HTML
 - Angular
 - Simpl Framework

DEMO

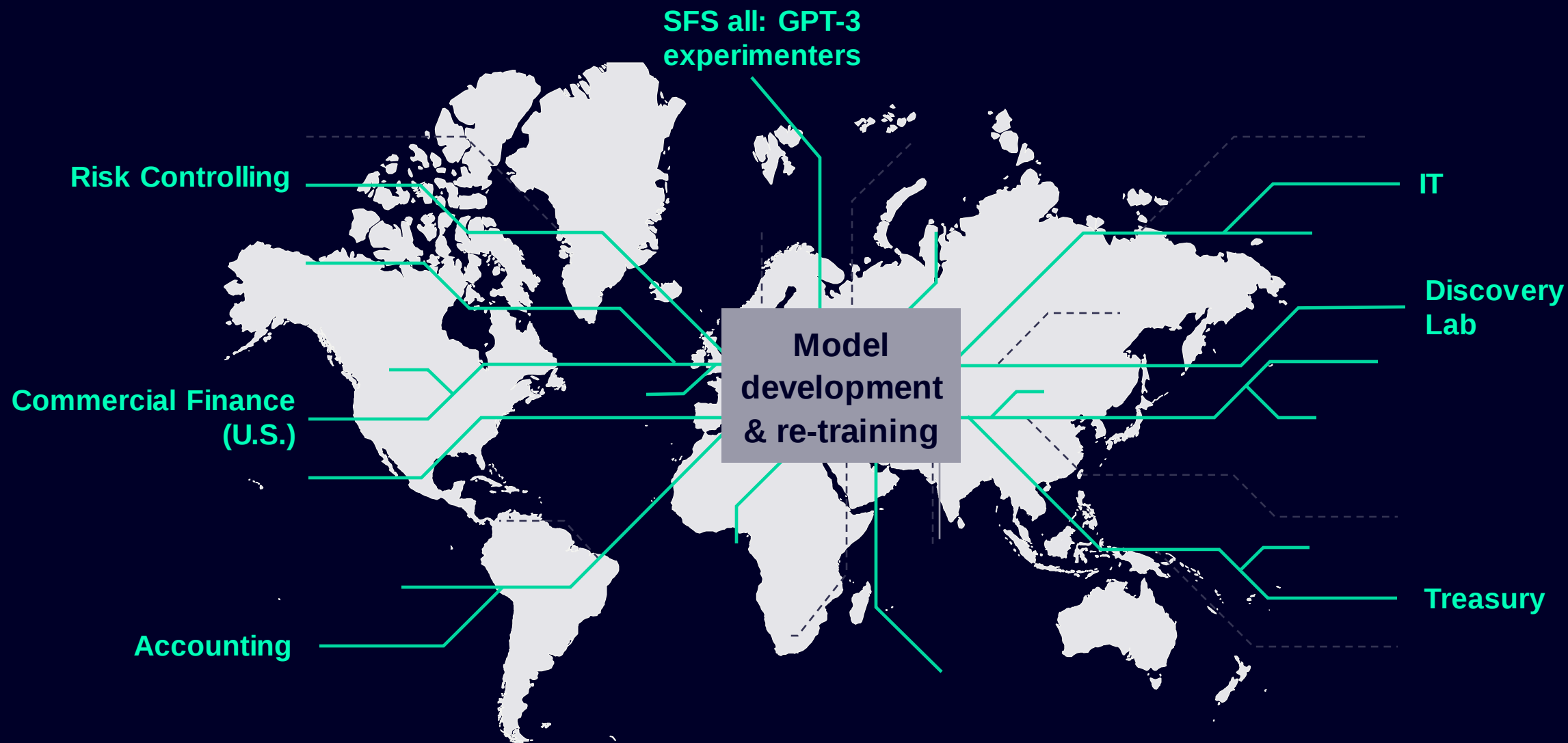
User perspective:

mark'd is used as an annotation platform for TEXT and PDF documents

- User uploads documents into mark'd
- The text is extracted from the documents and preprocessed for displaying it in the web browser
- Document data can be downloaded from the platform and processed in an NLP training pipeline
- Annotations can be created manually or upload from an AI algorithm



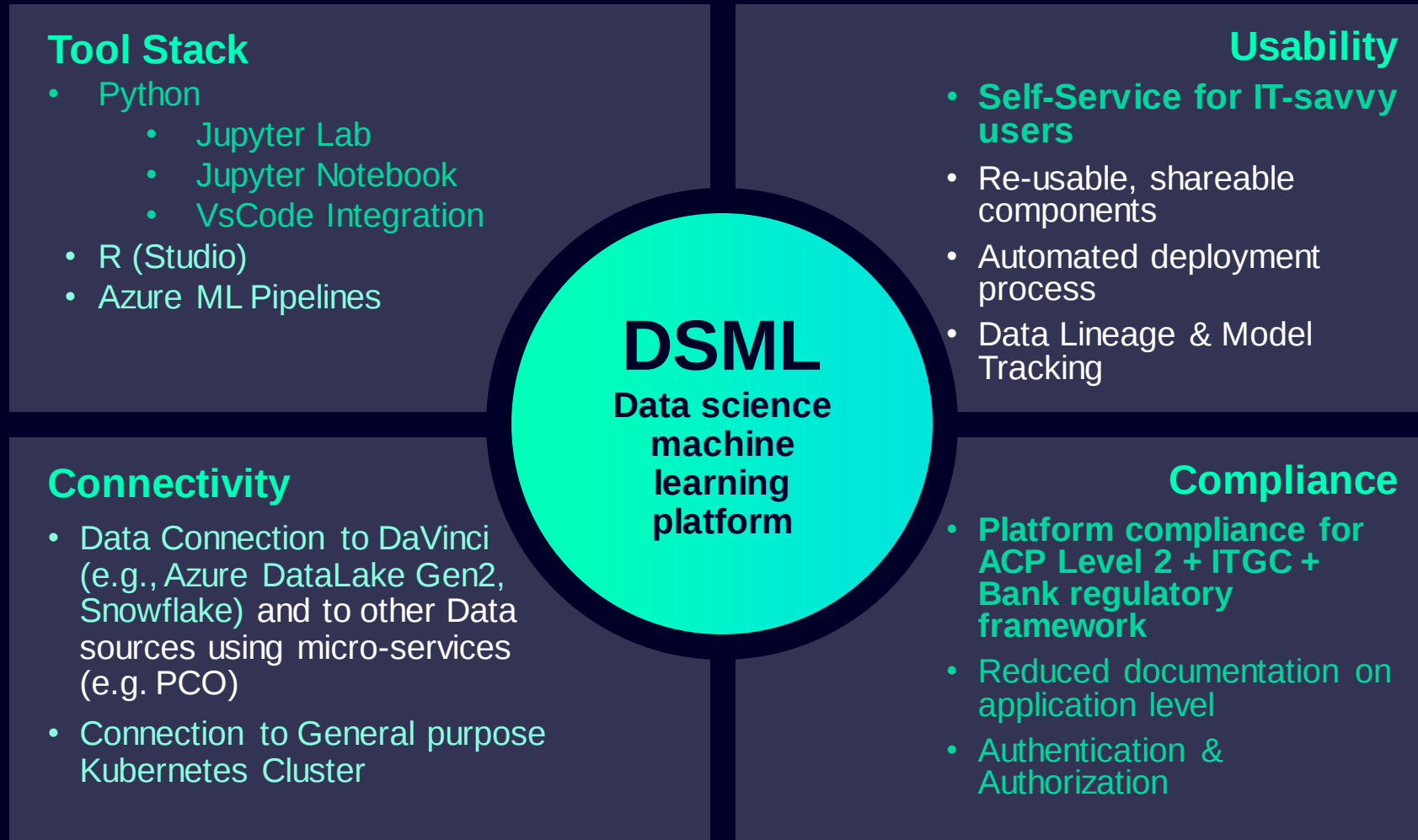
Data Science / Machine Learning Platform: user groups





DSML- (Data Science Machine Learning-) Platform

an end-to-end machine learning solution



Color code:

Existing

In the works

Coming up next



DSML (Data Science Machine Learning) Platform: usage

Applications using DSML today

PARDON (UK & Nordics)

→ RC

DALE (Digital Advanced Learning Emulator)

→ PS-EA

Prediction of early terminations

→ COF UK (done by Discovery Lab)

ADAP (Asset Disposition Advanced Predictions)

→ COF US

Applications coming up

MLDD (Machine Learning Due Diligence)

→ PS-EA

Outlier Detection

→ AC

Entity Mapping / Customer Mapping

→ RC CRI, potentially COF CRO RM

Guideline Search

→ PS-AM, other Bus and Functions

Chatbots based on GPT-3/openAI

→ SFS all



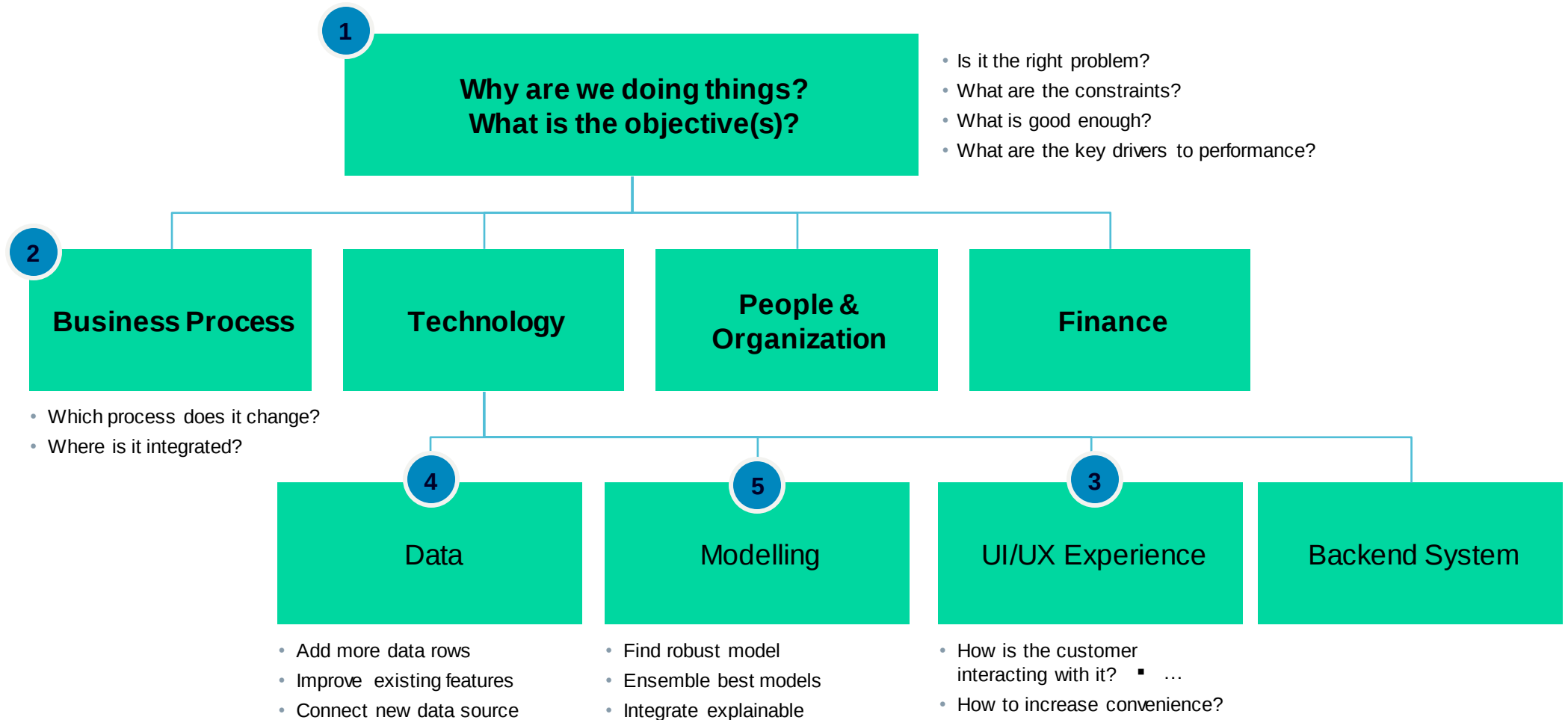
AI Journey: the building blocks

Start from the objective and work backwards to increase value creation

Why

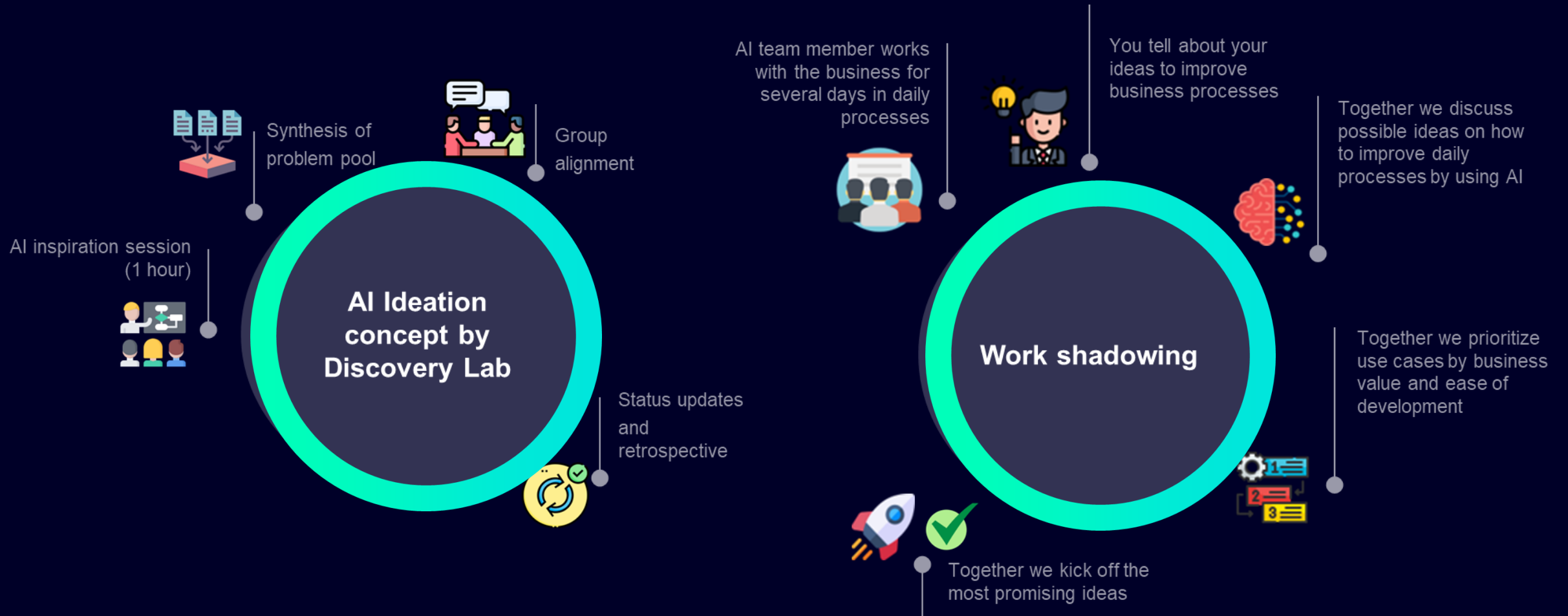
What

How

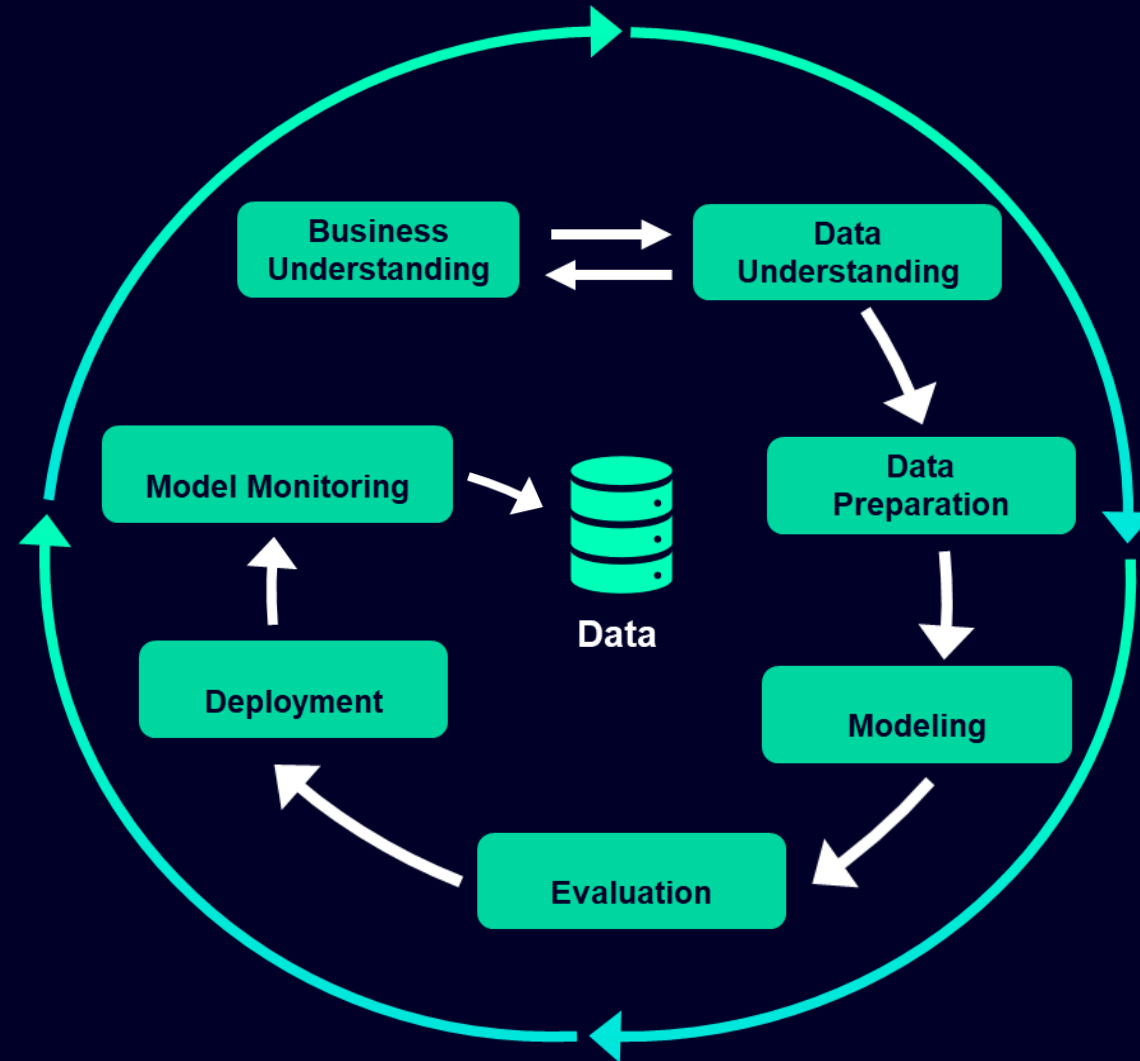


Adapted: Ian Xiao, Data Science is Boring, 2019; M. Taylor, E. Bernard, J. Gross, U. Waltinger, Assessing commercial scalability and return-on-invest with AI, AI with Purpose Summit, 2022

AI Journey: from Inspiration to Ideation to AI Project Kick-Off to tangible business results

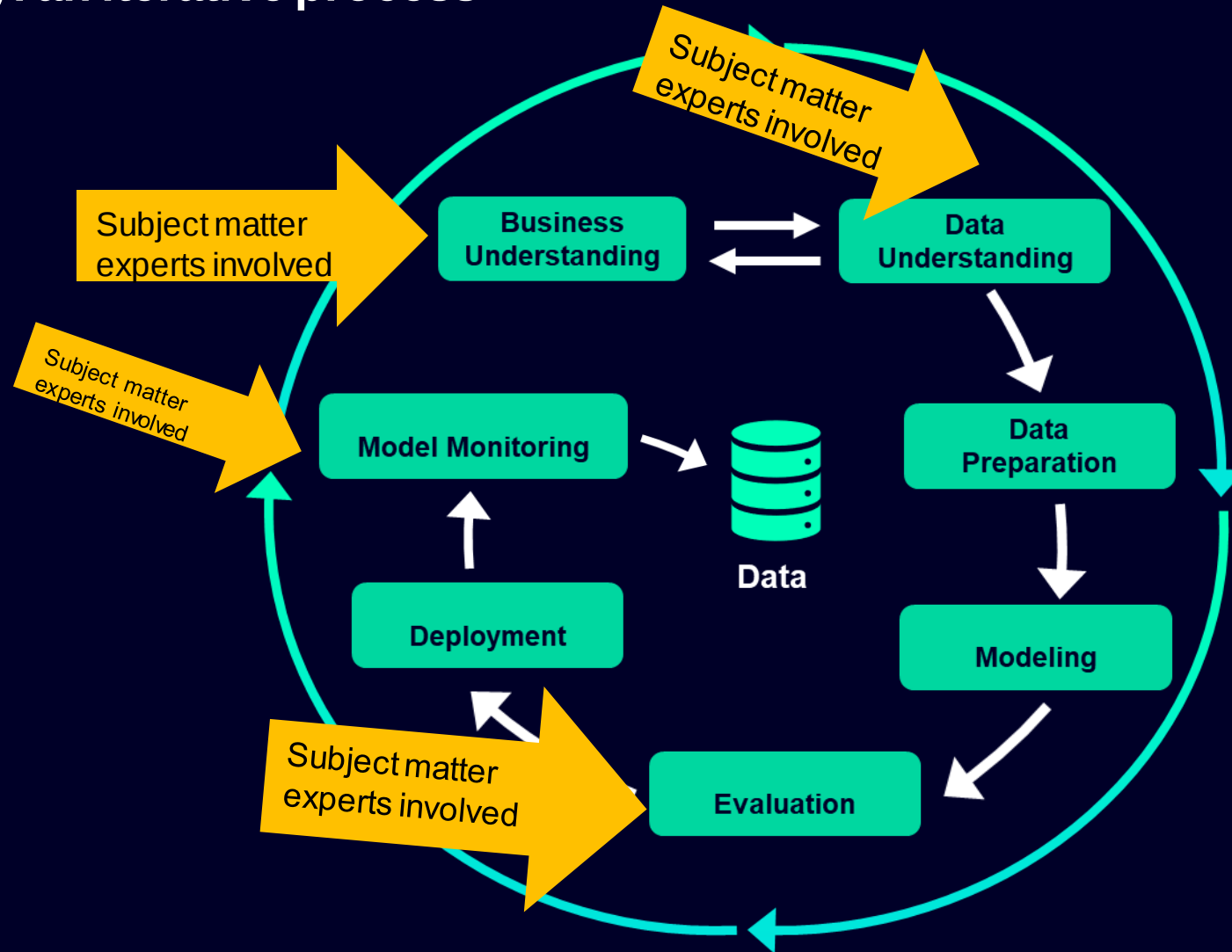


AI Journey: an iterative process



Source: Studer, S.; Bui, T.B.; Drescher, C.; Hanuschkin, A.; Winkler, L.; Peters, S.; Müller, K.-R.; Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology. Preprints 2021, 1, 0. <https://doi.org/>

AI Journey: an iterative process



Source: Studer, S.; Bui, T.B.; Drescher, C.; Hanuschkin, A.; Winkler, L.; Peters, S.; Müller, K.-R.; Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology. Preprints 2021, 1, 0. <https://doi.org/>



*Instead of focusing on the code, companies should focus on developing systematic engineering practices for improving data in ways that are reliable, efficient, and systematic. In other words, **companies need to move from a model-centric approach to a data-centric approach.***

Andrew Ng, CEO and Founder of LandingAI



| Contact

Published by Siemens SFS

SFS IT&CYS AI chapter

Otto-Hahn-Ring 6
81739, Munich
Germany

Phone +49 172 2562 135

E-mail it-ai.sfs@siemens.com

Ulrike.dowie@siemens.com



Back up

AI Use cases: examples

Output (Result to be predicted)	Input (Drivers, Impact Factors, Business Rules)	Application (expected solution)
Informational messages in chat, performance of tasks such as password resets	Data of calling customer, products services she uses, predefined actions based on her messages in chat	Help Desk Chat Bot
Whether the article is about Siemens or not	Text of the news article	News articles classification
Department to assign the ticket	Description of ticket and department that solved it	Ticket assignemt
Sentiment of the review (positive, neutral, negative)	Review in social media related to a product	Product Sentiment
Most similar document	Corpus of documents	Search engine
Groups of similar emails and topics per group	Customer emails describing technical problems	Document Segmentation
Correction of false custom tariffs, prediction of new tariffs	Description of products and their customs tariffs	Customs Tariffs
Similar interchangeable products	Product descriptions from various vendors / deparments	Product matching
Suitable jobs per CV, Suitable CVs per job	CVs, job ads	Job Matching

AI Use cases: examples

Output (Result to be predicted)	Input (Drivers, Impact Factors, Business Rules)	Application (expected solution)
Customer stays or leaves the company*	CRM, purchases, product usage, complaints, dunning, payment behavior	Customer churn prediction
Recommendation of product to customer	CRM, purchases or contracts, customer data, product data	Product recommendations (cross selling, up selling)
Customer pays back loan or not	Customer master data, her financial products, payment behavior, credit worthiness assessments	Credit evaluation
Lead is won or not	CRM Lead data, product, customer master data	Lead scoring
Fraudulent credit card transaction or not	Regular and fraudulent credit card transactions	Fraud detection
Product demand in the next months	Product demand in the past years, product characteristics, competitors, macroeconomics	Demand forecasting
Price of product	Product configuration (e.g. for car: cc, number of doors, age, km)	Price prediction
Revenue in the next month on BU, product and customer level	Past revenues, order intakes	Revenue forecasting



Horizontal AI Use Cases

Marketing

1 Brand insights

Understand in real time how customers are talking about your brand, identifying sentiment and extracting key topics and drivers.

NLP

2 Customer Insights

Identify customer segments, patterns and trends in customer interests by analyzing sources such as search engines, internal Siemens DI / SI / MO data, etc.

Pattern recognition

3 Feedback insights

Centralize the feedback channels in the company (email, social media, websites, customer portals etc.) and identify hot topics and dissatisfaction drivers.

NLP

4 Campaign optimization

Use AI to optimize the customers targeted or the content used, maximizing the conversion rates.

Pattern recognition

5 Retargeting

Identify the best content and next actions to reach customers who have engaged previously with another marketing activity.

Pattern recognition

Sales

1 Lead Scoring

Identify the marketing opportunities with higher sales probabilities to optimize the leads that are passed to sales and decide on the channel for each.

Classification

2 Pipeline Prioritization

Identify customers in the pipeline at risk of churn or likely to move the opportunity forward with an action.

Classification

3 Next-best action

Recommend the best potential next action by the seller for a specific account, given the past actions and pipeline state.

Recommendation

4 Meeting assistance

Provide guidance to the seller prior to a meeting or presentation to a customer, with context, opportunities, and recommended approaches.

Recommendation

5 Sales forecasting & quota setting

Predict the team's sales forecast on internal and external conditions for reporting purposes and quota setting.

Recommendation

Finance

1 Revenue forecasting

Complement traditional approaches to forecasting with AI-based methods based on historical sales data, information about deals in progress, and external conditions.

Regression

2 Contract management

Understand and mine contracts in the organization, identifying contract obligations and auditing contract terms.

NLP

Customer Service

1 Virtual agents

Resolve customer support issues with automated virtual agents that can interact naturally with conversational AI, or escalate to a customer agent if needed.

NLP

5 Issue identification

Quickly identify and responds to emerging trends in customer support and proactively address them.

NLP