

Skalierung von industrieller AI für Digital Manufacturing

Prof. Dr. Michael Heiss, Siemens AG Österreich

DI Roland Laucher, Siemens Digital Industries Software Austria & Switzerland

With an excellent PhD student and three years of time you can solve almost every challenge



...but how to scale?

Your examples of scaling AI applications excluding Large Language Models (LLMs)

Your examples of scaling AI applications

Image processing:



Scaling parameters

- Engineering: zero
- Customization: zero
- Applications domains: many
E.g. parking, toll, speed...
- SaaS offerings and complete solutions with camera
- Error rate close to zero
- If ticket is issued due to wrong detection – no safety issue: human as backup

license plate recognition system

→ Austria: W 66337 F



Does your AI Usecase have similar scaling properties?

Time-Series Models ☹️

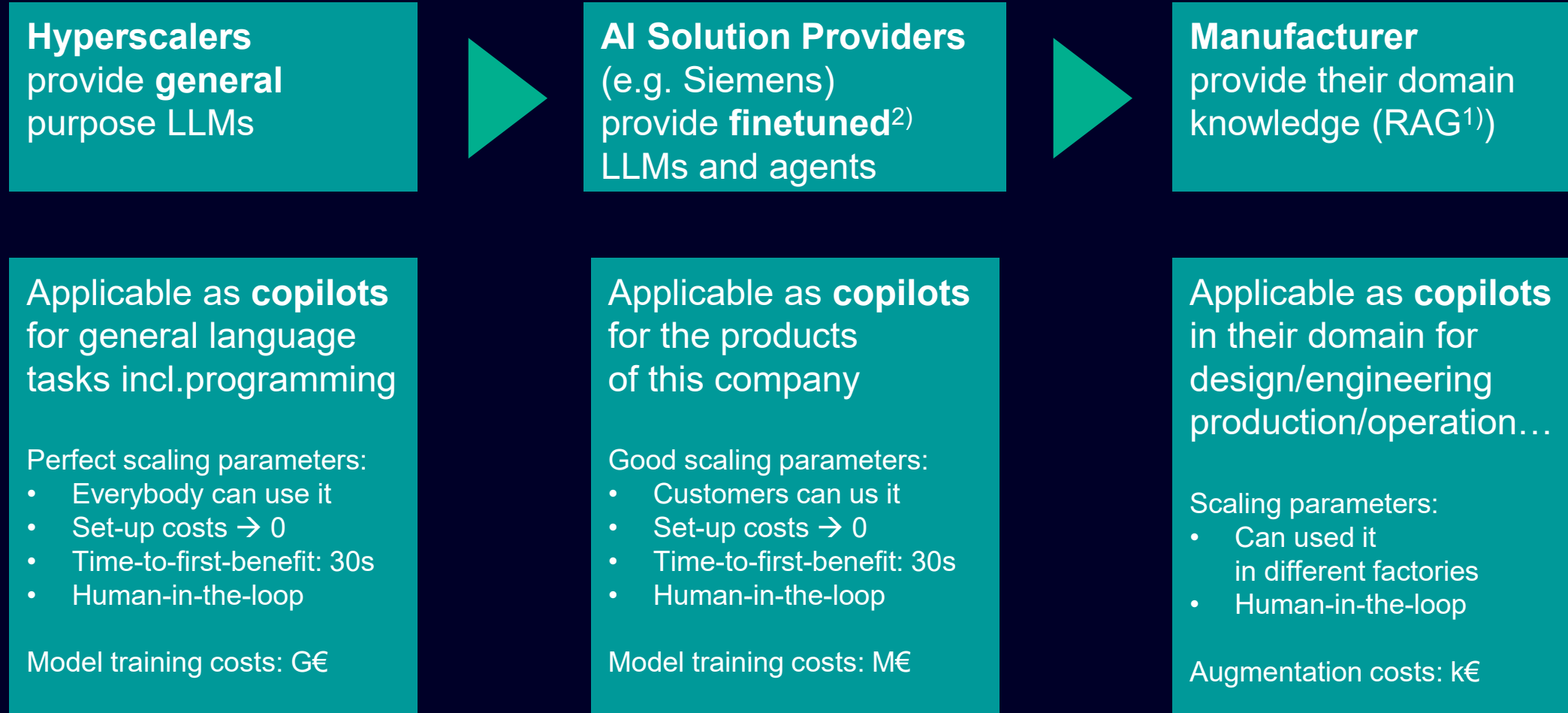
Image-Recognition Models 😐

Large Language Models 😊?

1 Scaling Strategies

2 Examples

Large Language Models (LLMs): a genius scaling approach for hyperscalers but the “last-mile-problem” remains in the industrial context



simplified

1) RAG: Retrieval-Augmented Generation

2) E.g. using intensive GPU-power from so-called AI-factories

**And how will
the time-series models
scale?**



Two scaling strategies for industrial AI (time-series...) : short term and long term

Blueprints

Already available
e.g.: Siemens with Amazon on AWS IDF¹⁾

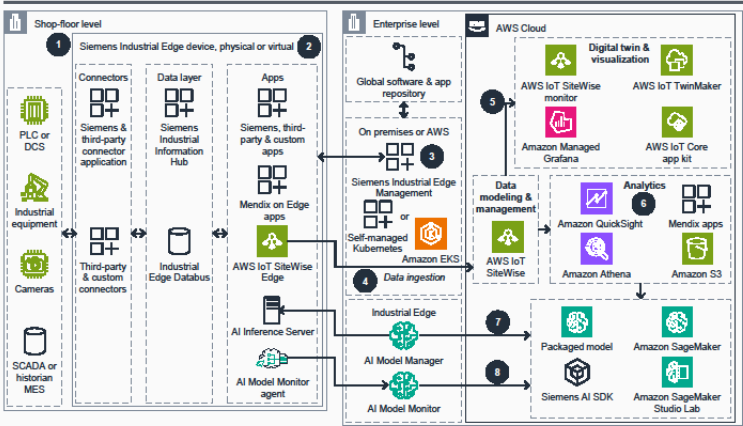
Industrial Foundation Model “Recharge Europe”



Work in progress

Guidance for Integrating an Industrial Data Fabric with Siemens Industrial Edge on AWS

This architecture diagram shows how to ingest near real-time data at scale from edge data sources into AWS IoT SiteWise by using AWS IoT SiteWise Edge and Siemens Industrial Edge.



- 1 Siemens Industrial Edge is an open software platform for simple, scalable, and manageable shop-floor IT. It provides decentralized and local data acquisition, storage, analytics, AI, and connectivity to AWS. The Industrial Edge Management application enables remote and central management of edge devices and applications.
- 2 On Industrial Edge devices, southbound connector applications—such as Open Platform Communications Unified Architecture (OPC-UA), Modbus TCP, and Siemens SIMATIC S7 connectors, ethernet and IP connectors, and other off-the-shelf connectors—collect data from industrial assets for on-premises processing and analysis. This is done with Siemens apps such as Energy Manager and Performance Insight. You can create your own apps using the Mendix on Edge integration, Industrial Edge Flow Creator, or Docker apps.
- 3 Industrial Edge Management can be deployed on-premises using a self-managed Kubernetes cluster or on AWS infrastructure using Amazon Elastic Kubernetes Service (Amazon EKS).
- 4 AWS IoT SiteWise Edge, deployed on Industrial Edge devices, collects and aggregates data and sends it to AWS IoT SiteWise.
- 5 AWS IoT SiteWise Monitor, AWS IoT TwinMaker, or Amazon Managed Grafana get data from AWS IoT SiteWise to create visualizations and get insights into collected industrial data.
- 6 Amazon Athena enables you to query cold Internet of Things (IoT) data from Amazon Simple Storage Service (Amazon S3) for data analytics with Amazon Managed Grafana, Amazon QuickSight, or Mendix low-code apps.
- 7 AWS artificial intelligence and machine learning (AI/ML) services, like Amazon SageMaker, use data from Amazon S3 to train ML models, then work with the Siemens AI Software Development Kit (SDK) to package and deploy ML models back to the edge.
- 8 The Siemens AI Model Manager deploys and manages ML models on the edge, and the Siemens AI Inference Server implements the models. The Siemens AI Model Monitor then observes them and provides results for use in model improvement.



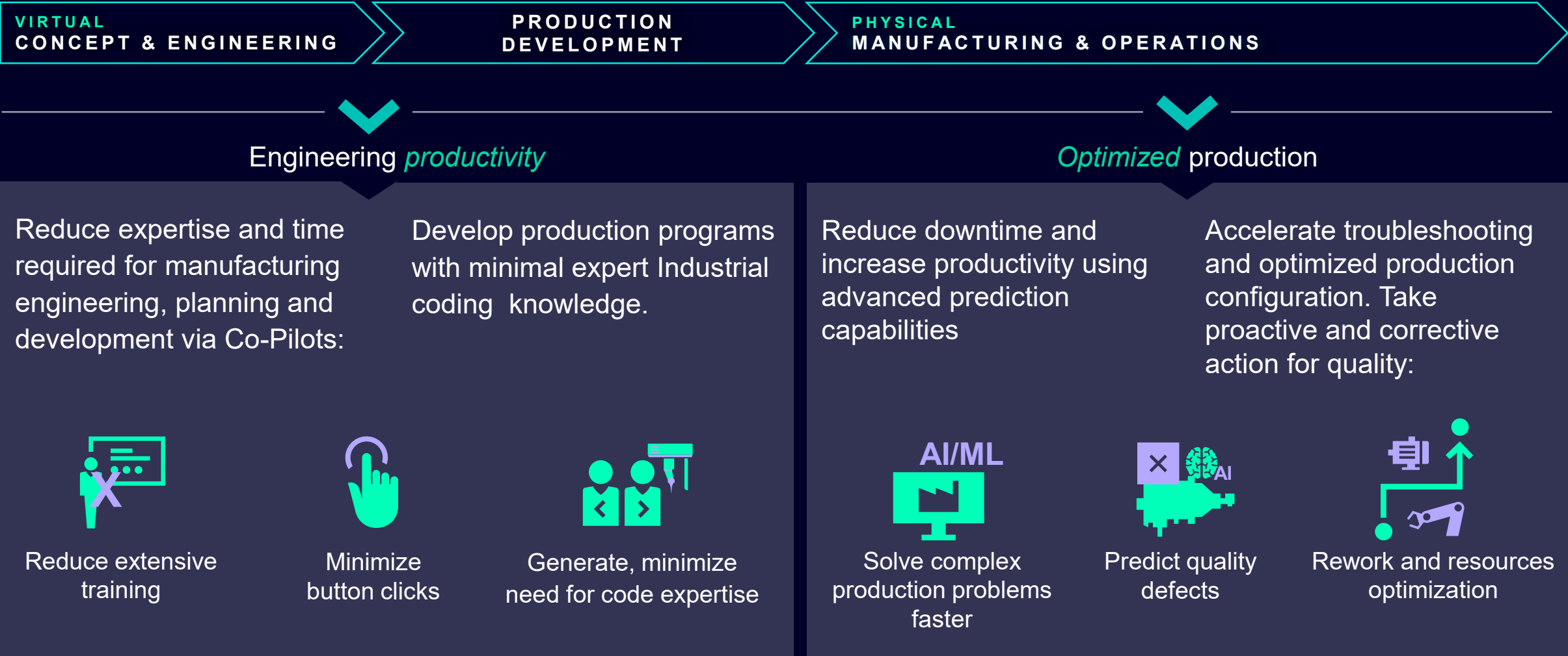
Industry data and context

1) IDF...Industrial Data Fabric

1 Scaling Strategies

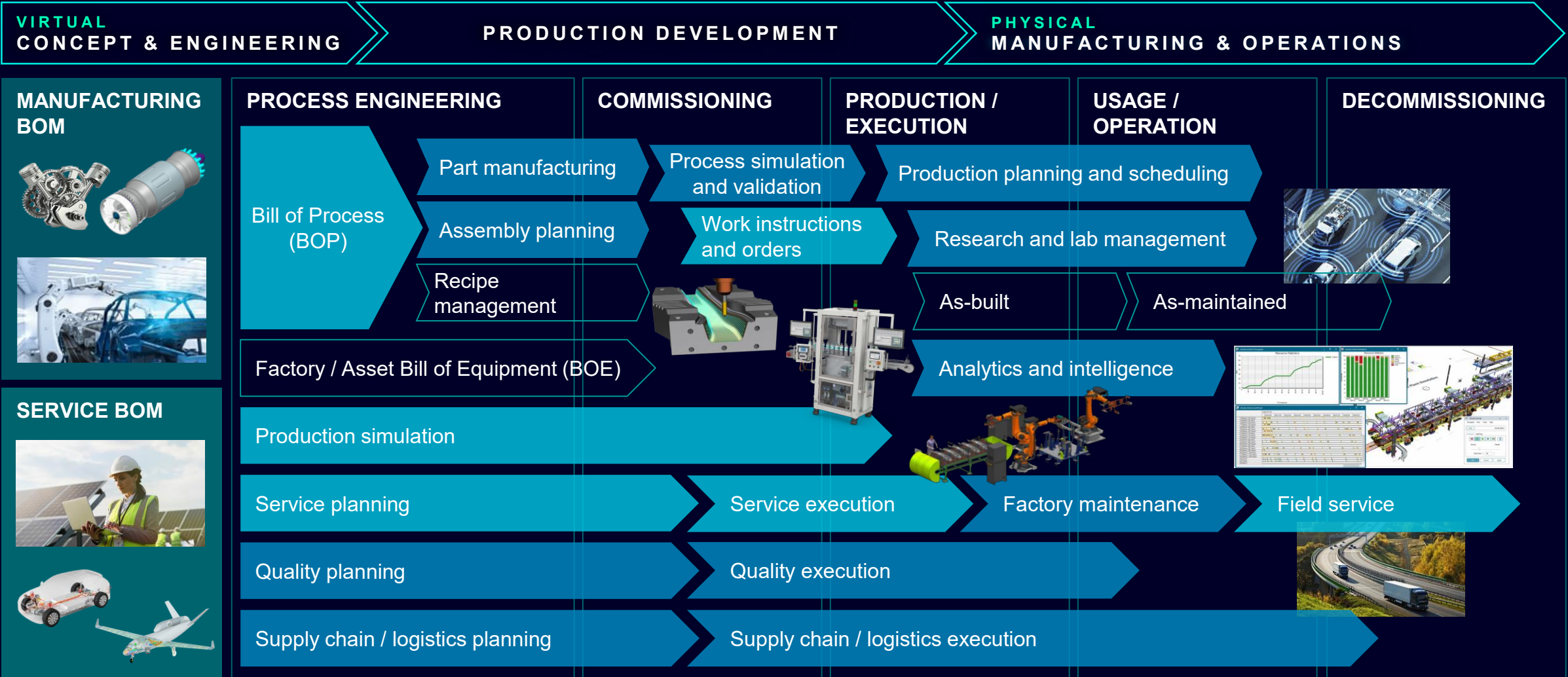
2 Examples

Benefits of AI in Digital Manufacturing



AI in Siemens Digital Manufacturing Portfolio

Ai



Tecnomatix Process Simulate Copilot

Challenge

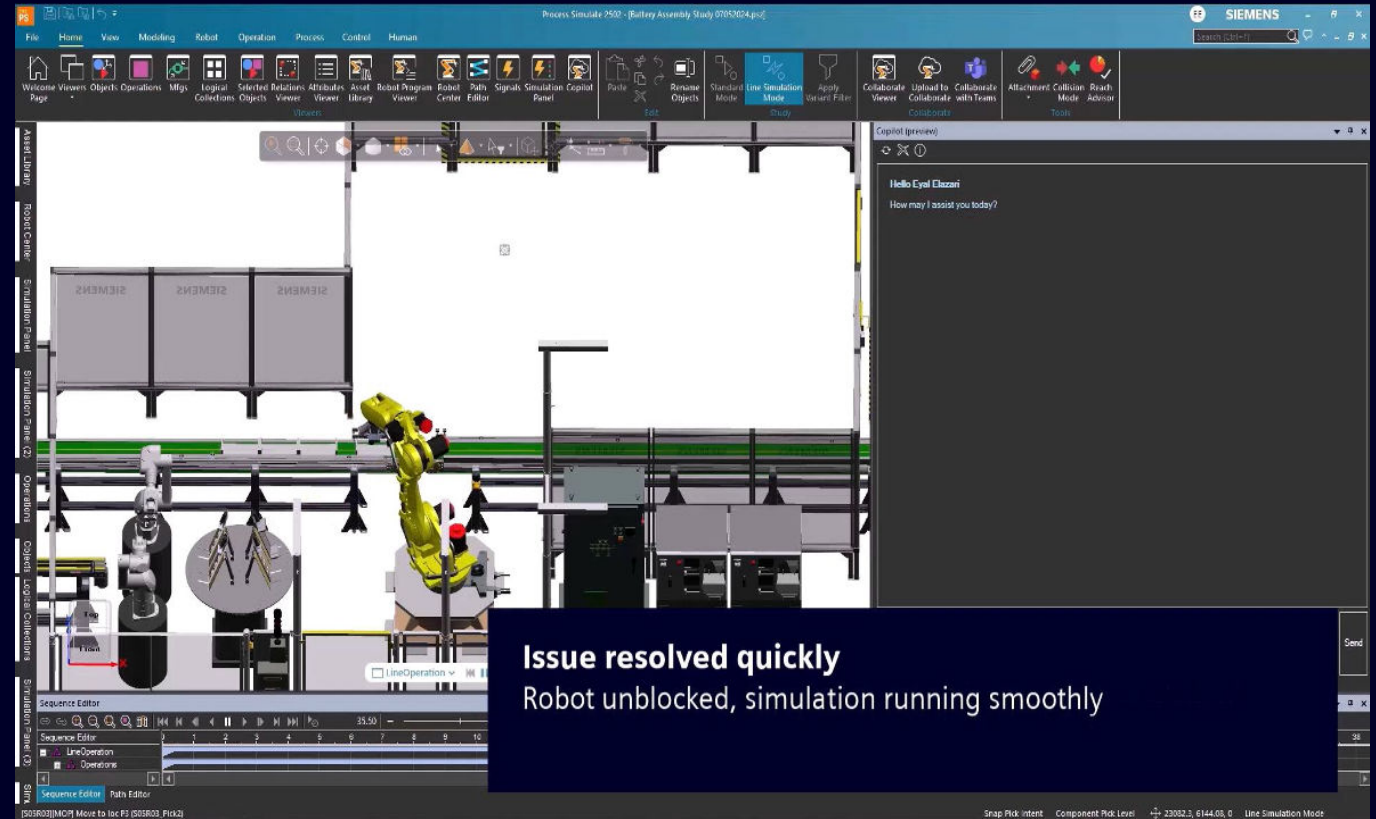
- **Optimized Efficiency:** Faster, more accurate troubleshooting in manufacturing.
- **Simplified Analysis:** Easier simulation study and data evaluation.

Solution

- Leveraging Generative AI (LLM), it analyses simulation data to provide insights, streamline troubleshooting, and suggest optimizations.

Benefit

- **Enhanced Efficiency:** Boosts productivity with quick, data-driven insights.
- **Optimized Operations:** Identifies and resolves issues for better performance.



Preview

Tecnomatix Process Simulate Collaborate Ergonomics

Challenge

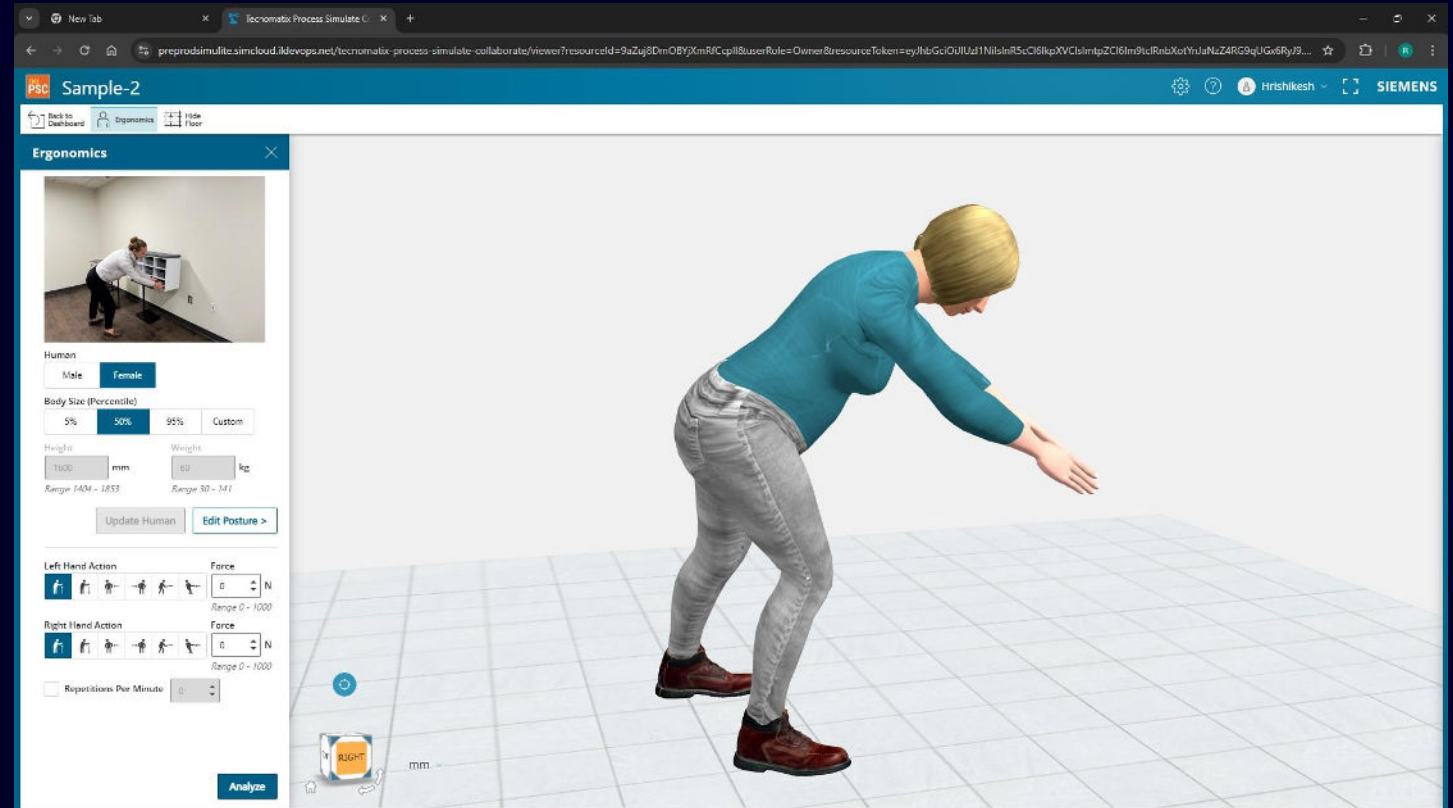
- Accurate 3D posture modelling is vital for valid ergonomic analyses but can be time-intensive & requires expertise.

Solution

- An AI-based tool to automatically generate 3d human poses from a basic photograph

Benefit

- Advanced posture modelling with minimal time and effort.
- Instant access to detailed results online.



[Link to Learn more](#)

Teamcenter Copilot

Challenge

- Searching for information
- Analyzing dense documents

Solution

- Teamcenter Copilot brings document intelligence, BOM exploration, and conversational search into a single, intuitive interface
- Fast, intelligent, context-aware assistance
- Secure, integrated, and grounded in your data

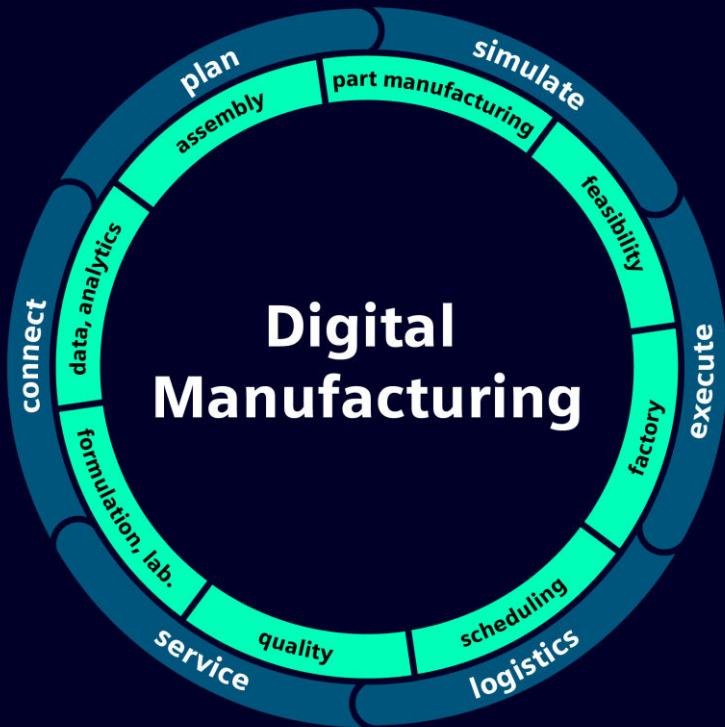
Benefit

- Improve engineering productivity

The screenshot displays the Siemens Teamcenter Copilot interface. The main window shows a document titled "STANDARD_Auto Requirements" with a preview of a table of contents. The table lists various requirements such as "Redundant Power Systems Requirements", "Dynamic Path Planning Systems Requirements", and "Electronic Stability Control (ESC) system". The right sidebar contains the "Copilot" panel, which provides a summary of the document and a table of diagnostic and safety requirements for ADAS. The table includes columns for Requirement, Description, and Specification.

Requirement	Description	Specification
Safe State Transition	Transition to safe state upon critical system failure	Within 200 ms of failure detection
Preemptive Safe State Activation	Detect early warning signs and activate safe state	Within 100 ms of identifying risks
Emergency Braking	Apply emergency braking during safe state transition	Within 100 ms of activation
Vehicle Stabilization	Maintain vehicle stability during safe state transition	Within 100 ms of failure detection
Driver Notification	Notify driver of safe state transition	Within 1 second of entering safe state
Controlled Path Deviation	Allow safe path deviation during transition	Not exceeding 1 meter , completed within 3 seconds
Normal Mode Reactivation	Return to normal driving mode after issue resolution	Within 10 seconds of issue resolution

Contact



Prof. DI. Dr. Michael Heiss
Siemens AG Österreich
Siemensstraße 90
A-1210 Wien | Austria

www.siemens.at/innovation-labs
michael.heiss@siemens.com
+43 664 8855 1526

<http://www.linkedin.com/in/michaelheiss>



DI. Roland Laucher
Portfolio Executive Digital Manufacturing
Siemens Digital Industries Software
Austria & Switzerland

www.sw.siemens.com
roland.laucher@siemens.com
+43 664 8863 4776

<https://www.linkedin.com/in/roland-laucher/>



SIEMENS