

# ERP 2025, Kufstein, Tirol

## Simulation und KI-generierte Vorhersage: Logistikprozesse planen

Prof. Dr. Carlo Simon

✉ [simon@hs-worms.de](mailto:simon@hs-worms.de)





## Es begann mit meiner Unfähigkeit ...

... direkt auf eine einfache Bemerkung zu reagieren:

Deine Simulation hilft mir nicht vorherzusagen,  
ob und wie sehr sich die Kunden in der nächsten Woche verspäten.

Und das ist offensichtlich wahr!

Habe ich einen echten Business Need übersehen? Was wäre eine gute Antwort auf diese Bemerkung gewesen?

Dafür muss ich von vorne beginnen!



# Die Simulation eines realen Fallbeispiels



(Infraserv Logistics GmbH, 2023)



# Fakten zur Fallstudie

## Basisfakten

- Platz für mehr als 21,000 Paletten
- 9 separate Lagerhaussegmente
- Lagermöglichkeit für verschiedene LGK Gefahrstoffklassen
- Lagertemperaturen von -6° C bis 20° C

(Infraserv Logistics GmbH, 2023)

## Mitarbeiter könnten Prozessvarianten simulieren,

- indem sie Zeiten und Ressourcen anpassen und
- den Schedule der Aufträge variieren.

## Ein Dashboard soll

- die Prozesse anschaulich visualisieren
- und aus verschiedenen Perspektiven betrachten helfen.



# Process-Simulation.Center (P-S.C) & Dashboard

## Process-Simulation.Center (P-S.C)

- Petri-Netz basiertes Integriertes Managementsystem (IMS)
- $\geq 500$  registrierte, zumeist akademische Anwender
- Höhere Petri-Netze werden als universelle Modellierungs- und Simulationsprache für dynamische Systeme eingesetzt
- Die Prozesssimulation wird durch die limitierenden Ressourcen bestimmt

## Dashboard

- Speziell für den Businesspartner entwickeltes Dashboard
- Verwaltung der Stammdaten
- Visualisierung der Lagertransporte im Tagesverlauf



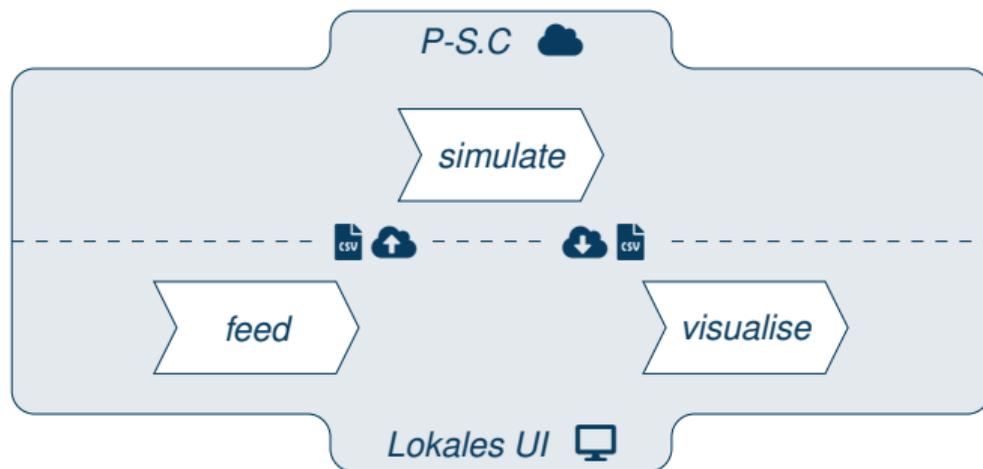
# Schritte von der realen Welt zum digitalen Abbild

Folgende **Schritte** wurden durchlaufen:

- 1 Modelliere die Phasen von Ein- und Auslagerungsprozess.
- 2 Spezifiziere die den Prozess kontrollierenden Daten.
- 3 Entwickle ein datengetriebenes Prozessmodell.
- 4 Entwickle ein UI für die Stammdatenpflege.
- 5 Entwickle ein UI zur Visualisierung des Prozessverlaufs.



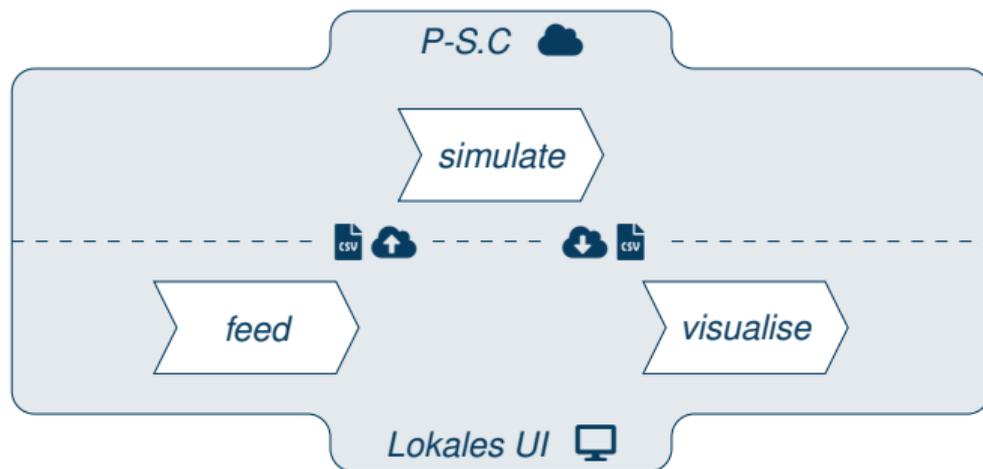
# Schritte von der realen Welt zum digitalen Abbild



Phase *feed*: Stammdaten für Ressourcen, Zeiten und Aufträgen als CSV Schnittstelle.



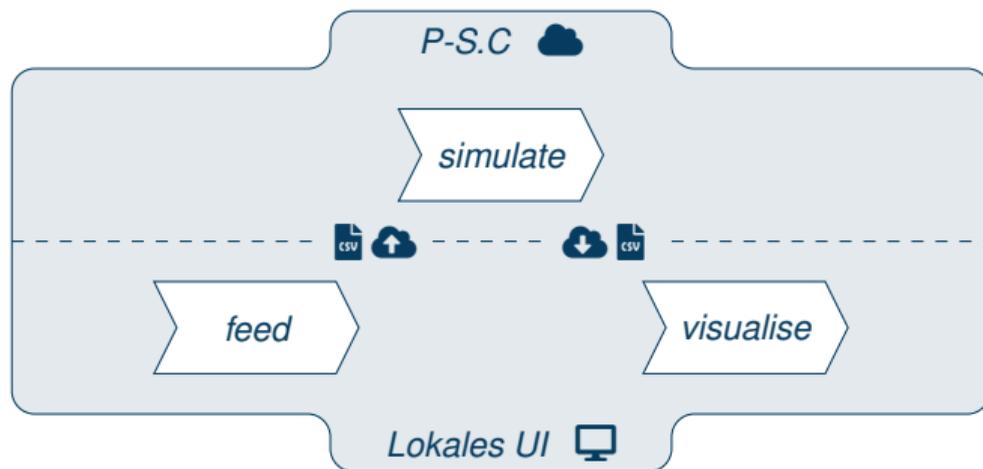
# Schritte von der realen Welt zum digitalen Abbild



Phase *simulate*: Das **P-S.C** simuliert daraufhin alle Lagerbewegungen und exportiert diese als CSV-Datei.



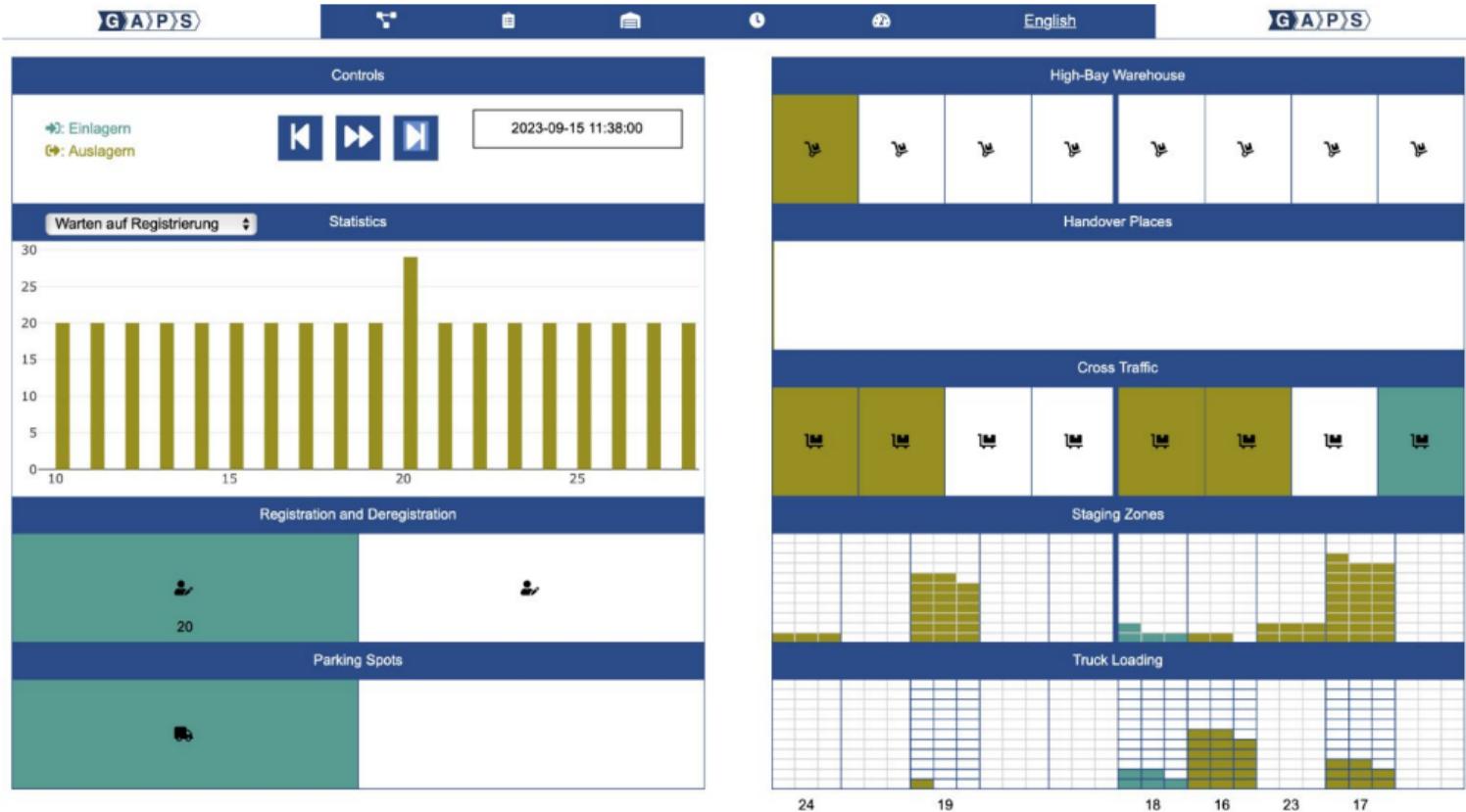
# Schritte von der realen Welt zum digitalen Abbild



Phase *visualize*: Visualisier die Lagerbewegungen auf Basis des Simulationsergebnisses.



# Ein Blick durchs Schlüsselloch





# Was kam dabei raus?

## Wir konnten ...

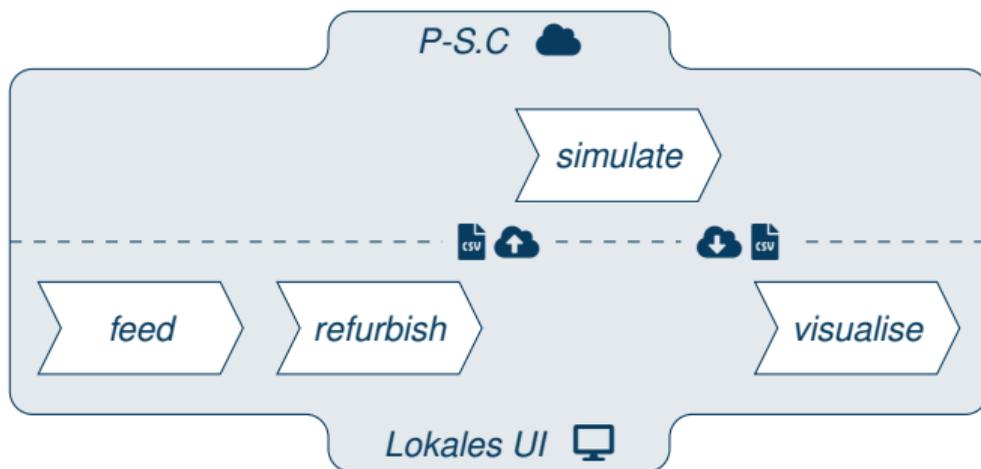
- die rückliegenden Engpässe erklären und
- das Werkzeug erlaubt Schedules hinsichtlich ihrer Toleranz gegenüber Veränderungen zu bewerten.

## Aber wir konnten nicht ...

- zu erwartende Veränderungen vorhersehen.
- Das können die Praktiker aber auch nicht, denn sonst wäre es nicht ihr Business Need gewesen.



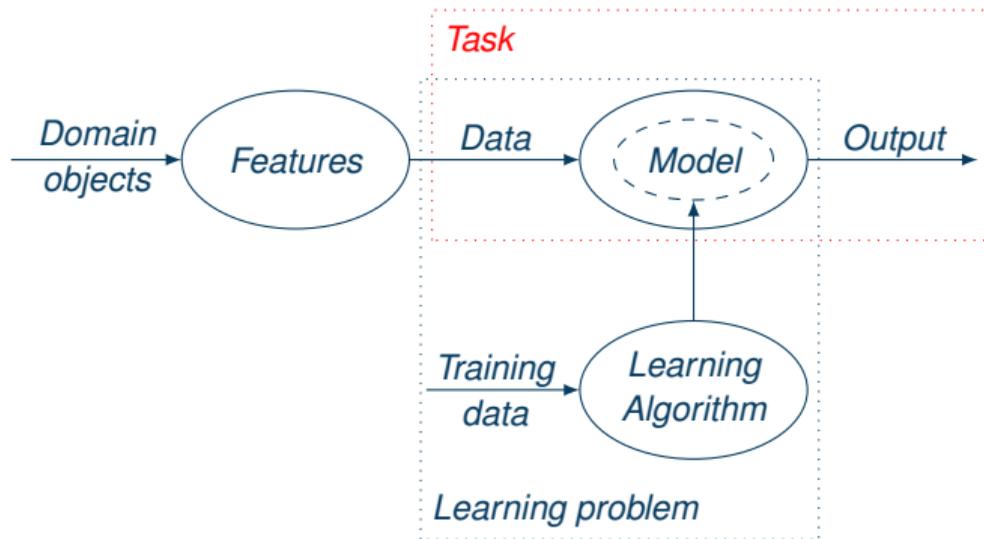
# Ein neuer Schritt



Phase *refurbish*: Kann Predictive AI (und insbesondere ML) die Daten für uns verbessern?



# Machine Learning



(Flach, 2012)



# Klassifikation vs. Regression

## Definition

- 1 Mathematisch gesehen ist eine Klassifikation eine Funktion  $y = f(x)$ , bei der  $x$  das gesicherte Wissen ist und  $y$  die Output-Kategorien.
- 2 Regression geht weiter und versucht die Datenpunkte zu identifizieren, die zu den Eingabedaten passen.

in Anlehnung an (Mattmann, 2020)

## Anwendung

- 1 **Klassifikation** könnte aufzeigen, welche Transporte unpünktlich sind.
- 2 **Regression** könnte den Grad der Verspätung z.B. in Minuten aufzeigen.



# Annahme und Voraussetzungen

## Annahme

Es gibt **Gründe** für Verspätungen. Diese mögen dem Logistikunternehmen unbekannt sein, aber sie lassen sich womöglich aus Vergangenheitsdaten ableiten.

## Voraussetzungen

- Wir brauchen Vergangenheitsdaten, um ein neuronales Netz zu trainieren und Vorhersagen abzuleiten.
- Wir müssen die **relevanten** Vergangenheitsdaten sammeln, in denen entsprechende Pattern enthalten sind.

## Beispiele

- Kommen manche Lieferanten typischerweise zu spät?
- Liefern manche Produzenten immer sehr früh?
- Sind Lieferungen an einem Mittwoch zuverlässiger als an einem Montag?



# Prototypische Implementierung

## Manchmal braucht man nur das richtige Buch!

- **NumPy** stellt Datentypen und Funktionen zur Verfügung, die den Umgang mit komplexen Strukturen wie Vektoren und Matrizen erleichtern.
- **pandas** ist für komplexere Strukturen und deren einfache Handhabung konzipiert. Eine Stärke ist die umfangreiche Funktionalität für Tabellenstrukturen.
- **Matplotlib** wird für visuelle Analysen und das Erstellen von Diagrammen verwendet.
- **scikit-learn** enthält viele ML-Algorithmen, die einfach im eigenen Programm genutzt werden können.
- **Keras** ermöglicht das Erstellen künstlicher neuronaler Netze.
- **TensorFlow** erweitert Keras um weitere leistungsstarke Funktionalitäten und kann große und komplexe Datenstrukturen verarbeiten.

(Karatas, 2024)



## Gegebene, unnütze und notwendige, aber fehlende Daten

Attribute	Description
<i>id</i>	order id
<i>product</i>	product group chem or pharma
<i>total</i>	total amount of pallets requested
<i>status</i>	initial or current order status
<i>ramp</i>	target ramp
<i>arrival</i>	scheduled time of arrival
<i>preparation</i>	scheduled time of completed staging
<i>fillHandover</i>	amount of pallets in handover areas
<i>fillRamp</i>	amount of pallets at target ramp
<i>fillTruck</i>	amount of pallets in truck
<i>usedGate</i>	used resource <i>gate</i>
<i>usedSGS</i>	used resource <i>SGS</i>
<i>usedVHS</i>	used resource <i>VHS</i>
<i>timestamp</i>	timestamp of the latest state change



# Gegebene, **unnütze** und notwendige, aber fehlende Daten

Attribute	Description
<i>id</i>	order id
<i>product</i>	product group chem or pharma
<i>total</i>	total amount of pallets requested
<i>status</i>	initial or current order status
<i>ramp</i>	target ramp
<i>arrival</i>	scheduled time of arrival
<i>preparation</i>	scheduled time of completed staging
<i>fillHandover</i>	amount of pallets in handover areas
<i>fillRamp</i>	amount of pallets at target ramp
<i>fillTruck</i>	amount of pallets in truck
<i>usedGate</i>	used resource <i>gate</i>
<i>usedSGS</i>	used resource <i>SGS</i>
<i>usedVHS</i>	used resource <i>VHS</i>
<i>timestamp</i>	timestamp of the latest state change



## Gegebene, **unnütze** und notwendige, aber fehlende Daten

Attribute	Description
<i>total</i>	total amount of pallets requested
<i>status</i>	initial or current order status
<i>arrival</i>	scheduled time of arrival



## Gegebene, unnütze und **notwendige, aber fehlende** Daten

Attribute	Description
<i>total</i>	total amount of pallets requested
<i>status</i>	status (inbound or outbound)
<i>arrival</i>	scheduled time of arrival (broken up)
<i>delay</i>	actual arrival time or delay
<i>distance*</i>	inside or outside the industry park
<i>agent</i>	shipping agent
<i>producer</i>	of the good

\* kann durchaus schwierig vorab zu bestimmen sein!



# Klappt das wirklich?

- Der bisherige Industriepartner sammelt diese Informationen derzeit nicht!
- Gleiches gilt für (durchaus große) Logistikunternehmen in Worms.

Aber dann sind wir auf den Beitrag von Constante-Nicolalde et al. (2020) gestossen!

- In dem Paper geht es eigentlich um Fraud-Detection, aber
- der Datensatz birgt noch einen anderen Schatz.

Der von den Autoren genutzten Datensatz steht frei zur Verfügung.

siehe (Constante et al., 2019)



## Zum Datensatz

- Der Datensatz umfasst mehr als 170.000 Handelsdaten, die von DataCo Global zur Analyse der Fraud-Detection genutzt wurde.
- Neben Details zum gelieferten Produkt, zum Lieferziel etc. enthält der Datensatz die geplante und tatsächliche Lieferdauer.
- 80% der Daten wurden nun als Trainingsgrundlage verwendet.
- Mit den restlichen 20% wurde eine Regression der Verspätungen vorgenommen.

*Für ca. 90% der Lieferungen  
konnte die Verspätung sehr präzise vorhergesagt werden!*

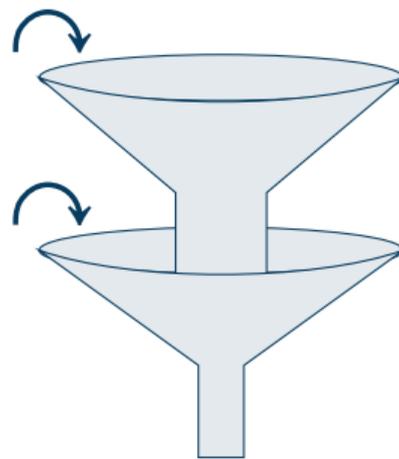
Da die Auslieferungen nicht instantan mit der Bestellung erfolgt sind, ergibt sich hier ein Optimierungspotential.



# Wie soll dieses Optimierungspotential gehoben werden?

*Tatsächliche  
Bestellungen*

*ML korrigierte  
Bestellungen*



*Optimierte  
Auslieferzeitpunkte*

- Mögliches Optimierungspotential bestimmen.
- Auf dieser Basis potentielle Industriepartner gewinnen und Ansatz weiter testen.
- Gegebenenfalls Integration ins P-S.C.

# Interesse? Kontaktieren Sie uns!

.... besuchen Sie die Webseite der **Group for Applied Processsimulation (GAPS)**

<https://www.hs-worms.de/en/gaps/>



... und nehmen Sie gerne mit uns Kontakt auf.

- Constante, F.-V.; F. Silva und A. Pereira (2019): DataCo smart supply chain for big data analysis. <https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis> (last accessed 15.05.2025).
- Constante-Nicolalde, F.-V.; P. Guerra-Terán und J.-L. Pérez-Medina (2020): Fraud Prediction in Smart Supply Chains Using Machine Learning Techniques. In: *Applied Technologies: First International Conference, ICAT 2019*, Hg. M. Botto-Tobar; M. Zambrano Vizuete; P. Torres-Carrión; S. Montes León; G. Pizarro Vásquez und B. Durakovic. Springer, Quito, Ecuador, S. 145–159.
- Flach, Peter (2012): *Machine Learning - The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press, Cambridge, 9. Aufl.
- Infraserv Logistics GmbH (2023): Overview hazardous substances warehouse. <https://www.infraserv-logistics.com/en/is1/news/news/> (last accessed 15.08.2023).
- Karatas, M. (2024): *Development of AI applications (in German: Eigene KI-Anwendungen programmieren)*. Rheinwerk Computing, Bonn.
- Mattmann, C., Hg. (2020): *Machine Learning with TensorFlow*. Manning, Shelter Island, NY, 2. Aufl.