From Prototype to Operative Software with RapidMiner

Data Analytics @ Lufthansa

Agenda

1 Lufthansa Industry Solutions: Who We Are

2 Our Daily Challenges

3 FlightPrediction @ Lufthansa

4 Lufthansa Industry Solutions & RapidMiner

We are

Lufthansa Industry Solutions

IT Consulting and Systems Integration

Lufthansa Industry Solutions combines the dynamics of an SME with the economically secure background of Lufthansa; an internationally acting global corporation.

Services overview

Lufthansa Industry Solutions

Strategy and Management Consulting			IT Service Provider		
Strategy Consulting	Organization & Process Consulting		IT System Integration & Development	IT System Operations	Operations of full business processes
	Organi-sational structure	Process Consulting / Process Organization		Application Mgmt. & Operations / Infra-structure	BPO specialized provider
		IT Consulting (processes, technologies, infrastructure)			

IT Consulting and System Integration

Lufthansa Industry Solutions

Our

Locations

Lufthansa Industry Solutions

Our

Locations

Hard Facts

Managing Director
Bernd Appel

Founding
1998 as Industry Solutions division within Lufthansa Systems AG

100% owned by
Lufthansa

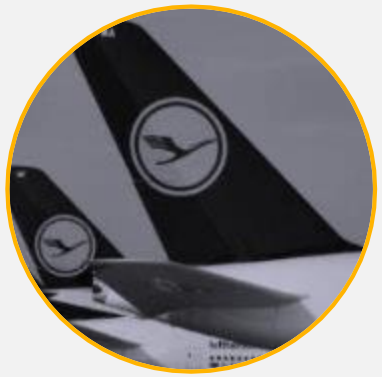
208 Mil. €
total revenue (46% of which within the Lufthansa Group)

>1300
trained employees

>200
customers

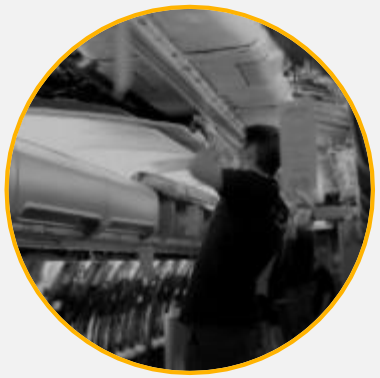
Lufthansa Industry Solutions

Lufthansa



Passenger Airlines

- Lufthansa Passenger
- Swiss
- Austrian

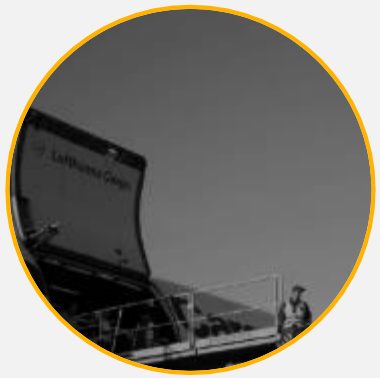


Aviation Industry

- Lufthansa Technik
- LSG Sky Chefs

Finance and Service Companies

- AirPlus
- LGBS



Air Freight

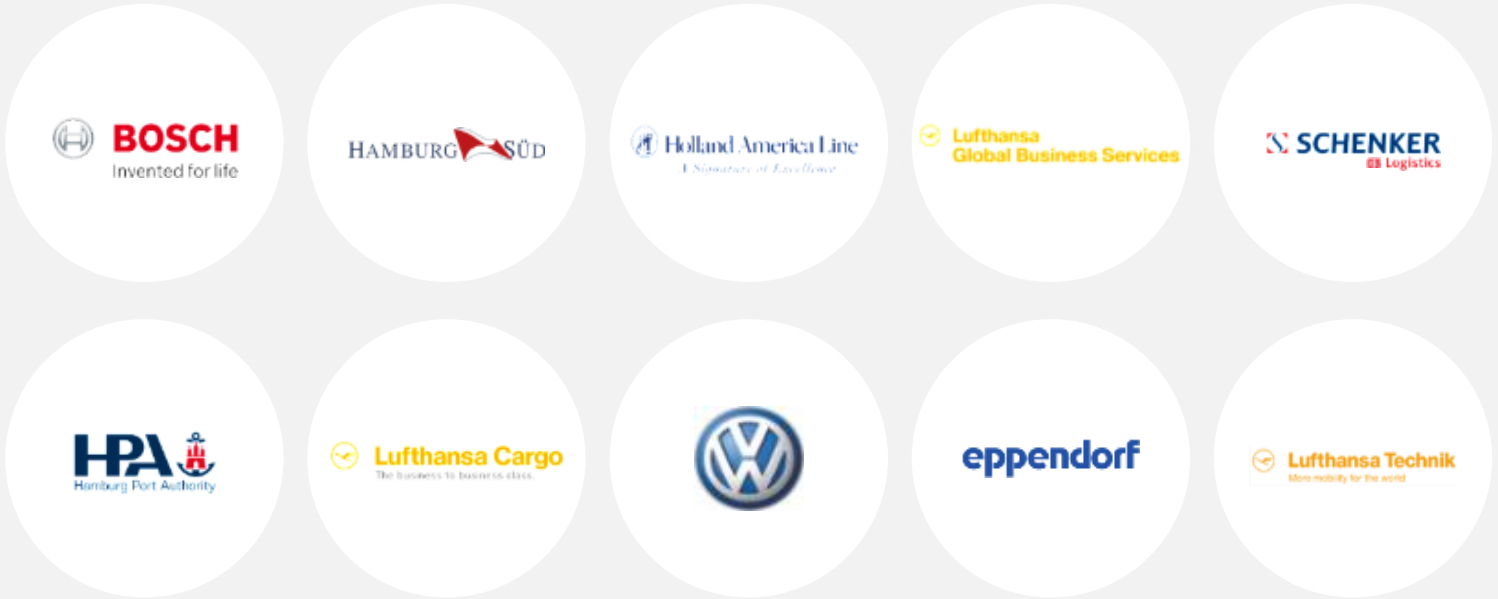
- Lufthansa Cargo

Digital Service Companies

- **Lufthansa Industry Solutions**
- Lufthansa Systems

Many

Projects Successfully Delivered



Lufthansa Industry Solutions

About Me

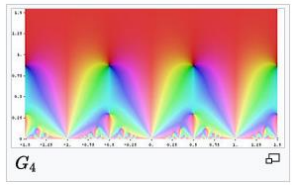
Quick Facts

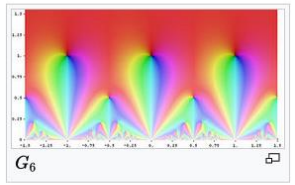
Fabian Werner

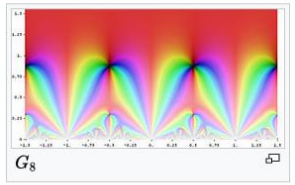
Born 1985

Ph.D. Mathematics (Number theory)

Past







$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{m=1}^{\infty}\sigma_{k-1}(m)e^{2\pi i m\tau}$$

Present

- Predictive analytics:
- Classification, regression, recommender systems, probabilistic modelling, …
- RapidMiner, R, Python, PHP, Teradata, Oracle, Hadoop, …

Future

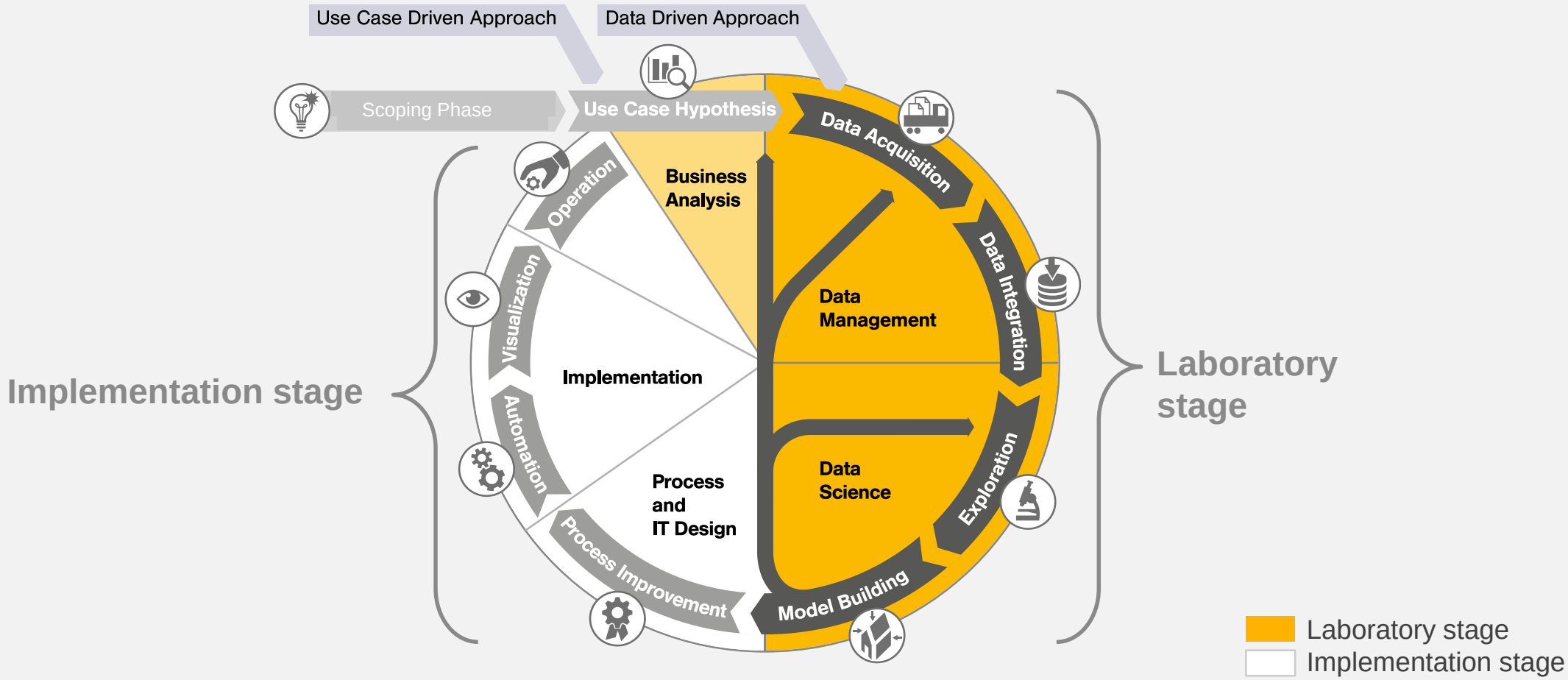
- Probabilistic modelling
- Time series analysis
- Bayesian learning
- Learning theory

Our Daily Challenges

Lufthansa Internal

Dashboarding	Predictive Analytics	(Deep) Machine Learning	Data Ingestion
Lufthansa	Lufthansa Cargo		Lufthansa Technik
▪ Airplane punctuality ▪ Competitor information ▪ Airfreight load monitor	▪ Arrival times ▪ Passenger connection ▪ Fuel & Weight forecast ▪ Predictive maintenance ▪ Recommend fligths	▪ Document and image recognition ▪ Damage recognition via audio data ▪ Bots (see Mildred)	▪ Live flight data ▪ Weather data streams ▪ Airplane sensor data

Process Model

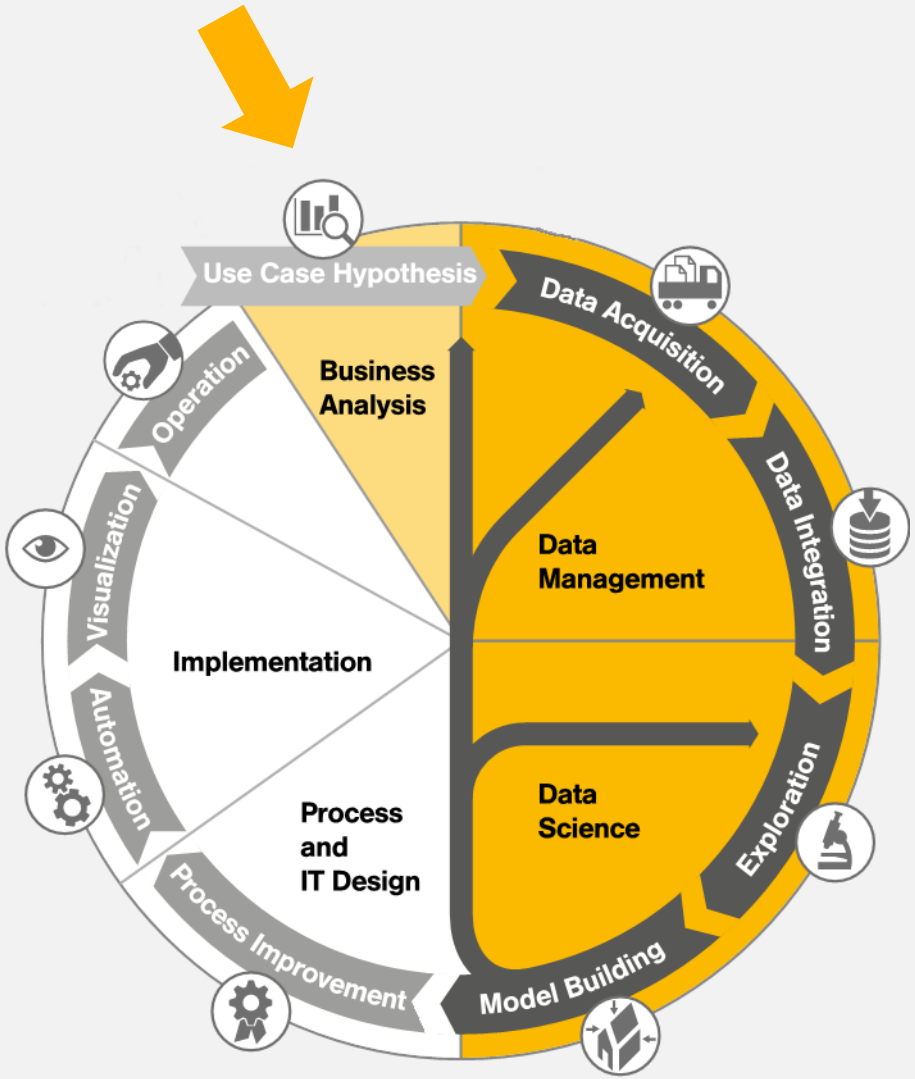


Lufthansa Industry Solutions

FlightPrediction @ Lufthansa

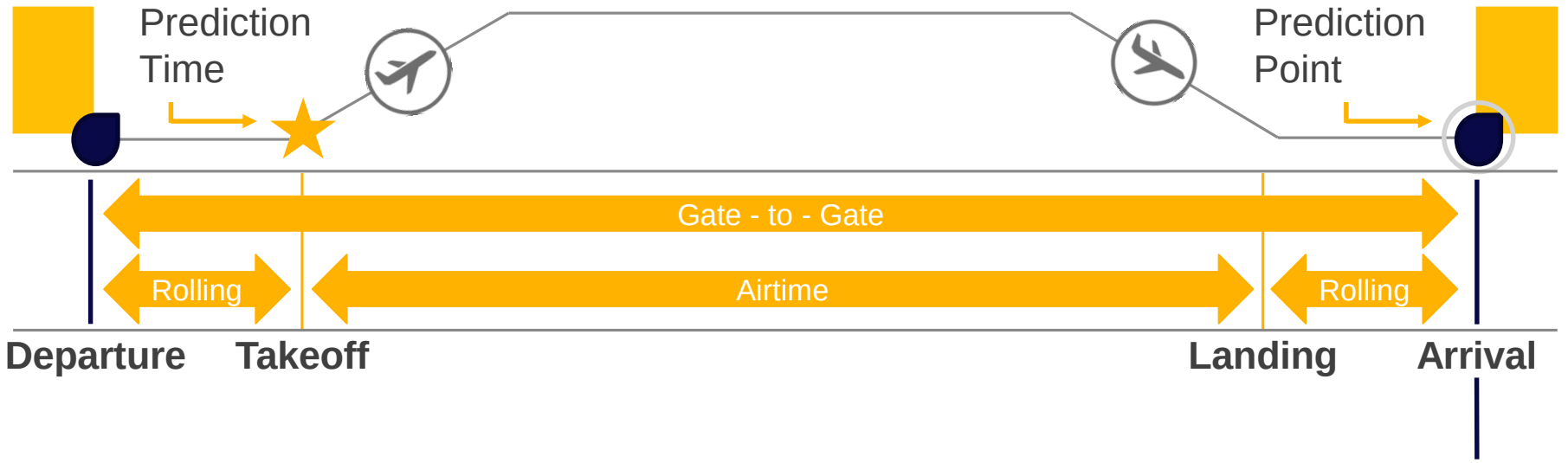
We predict the arrival time on aircraft takeoff

The Cycle

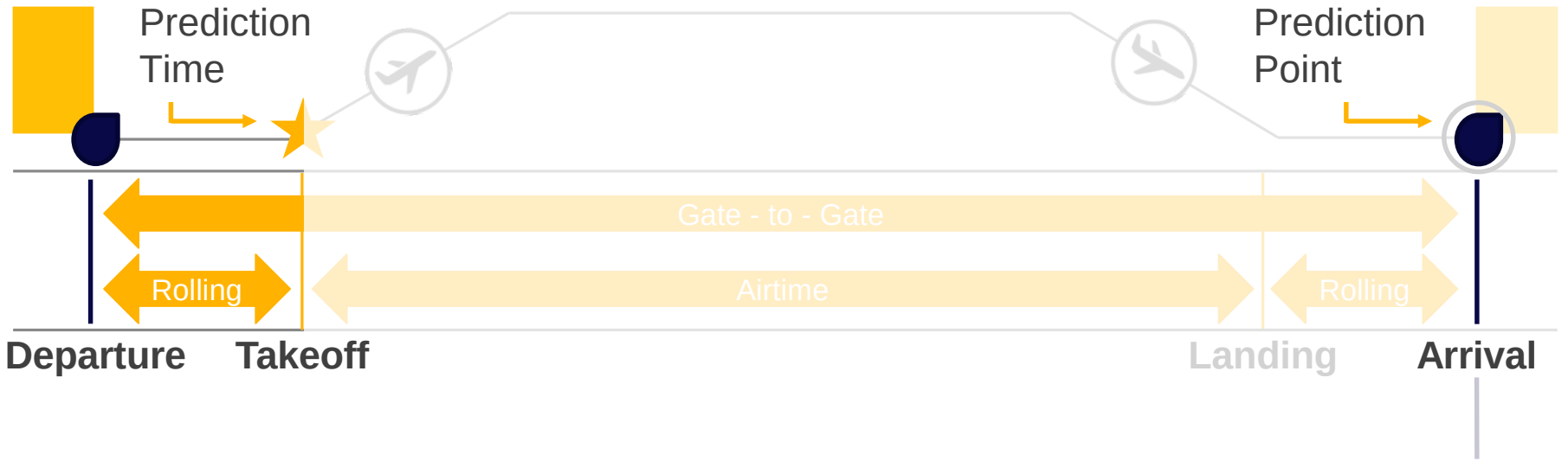


Lufthansa Industry Solutions

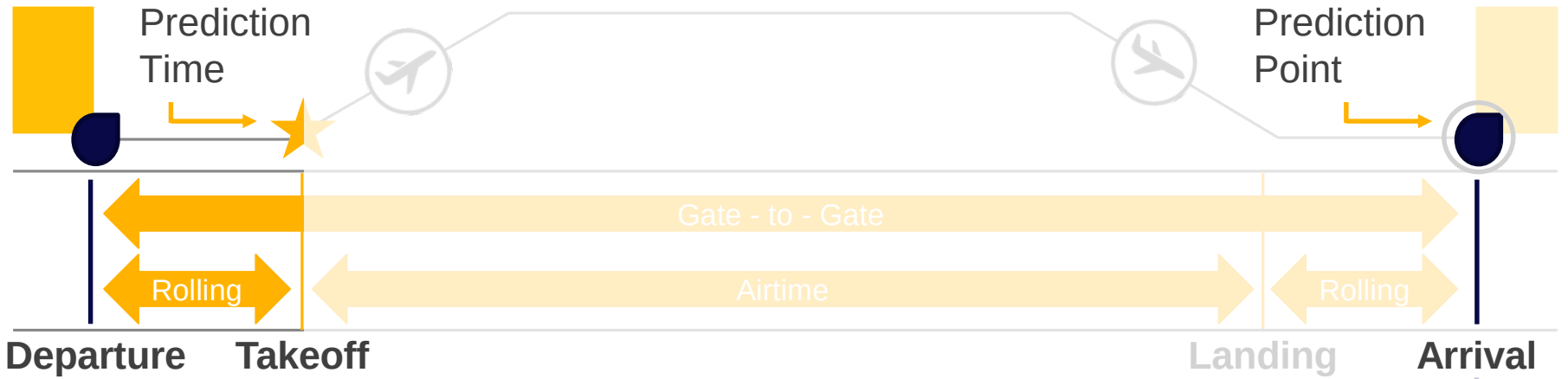
Goal & Process

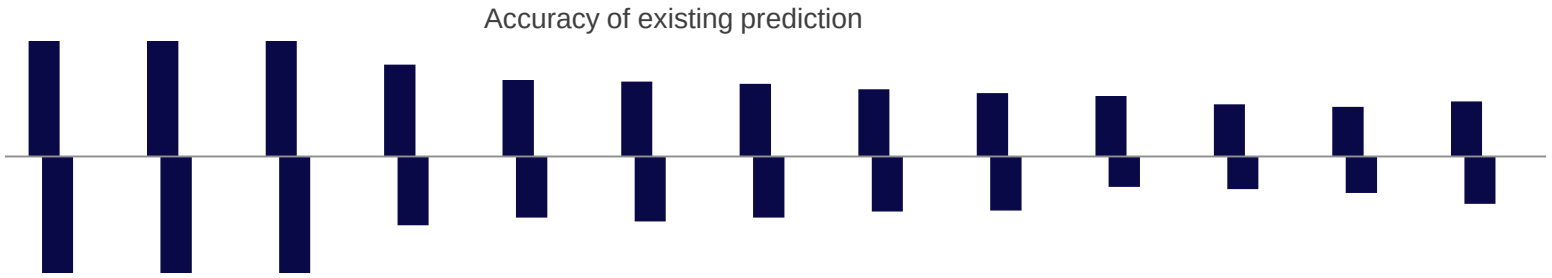


Goal & Process



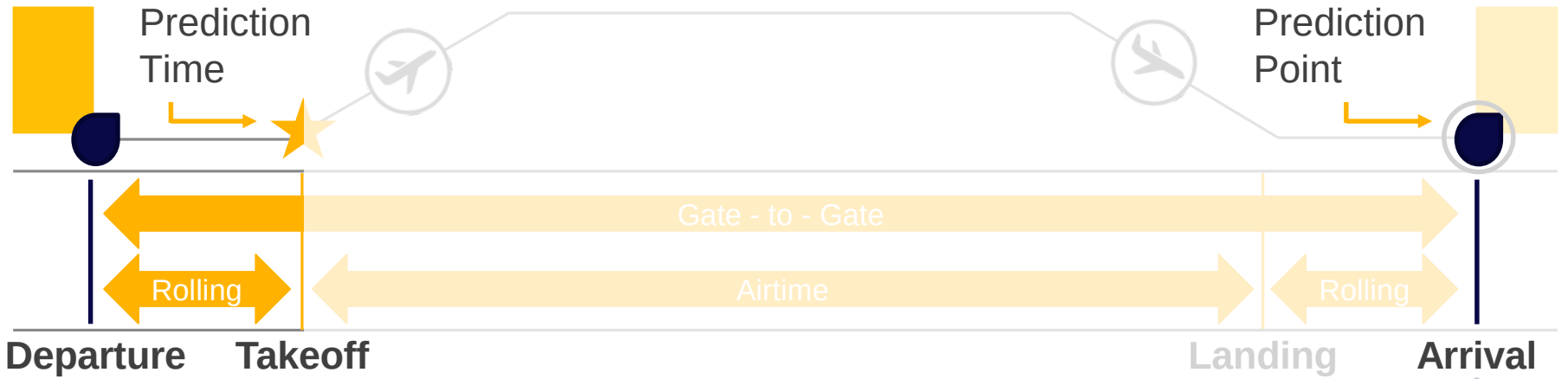
Goal & Process



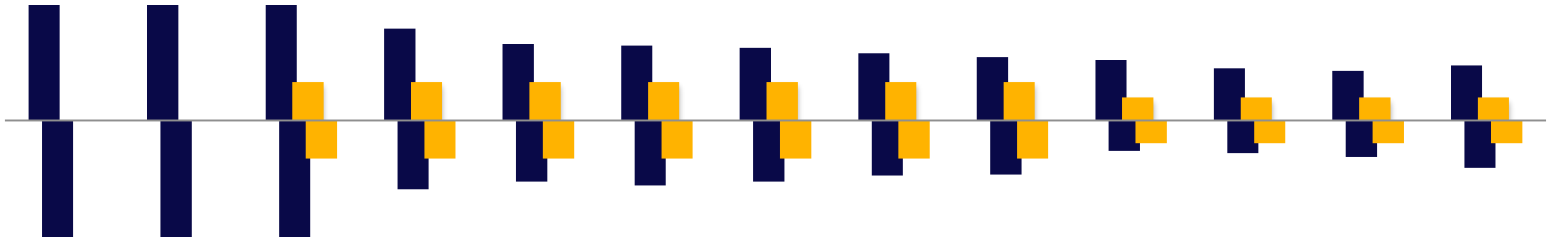


Lufthansa Industry Solutions

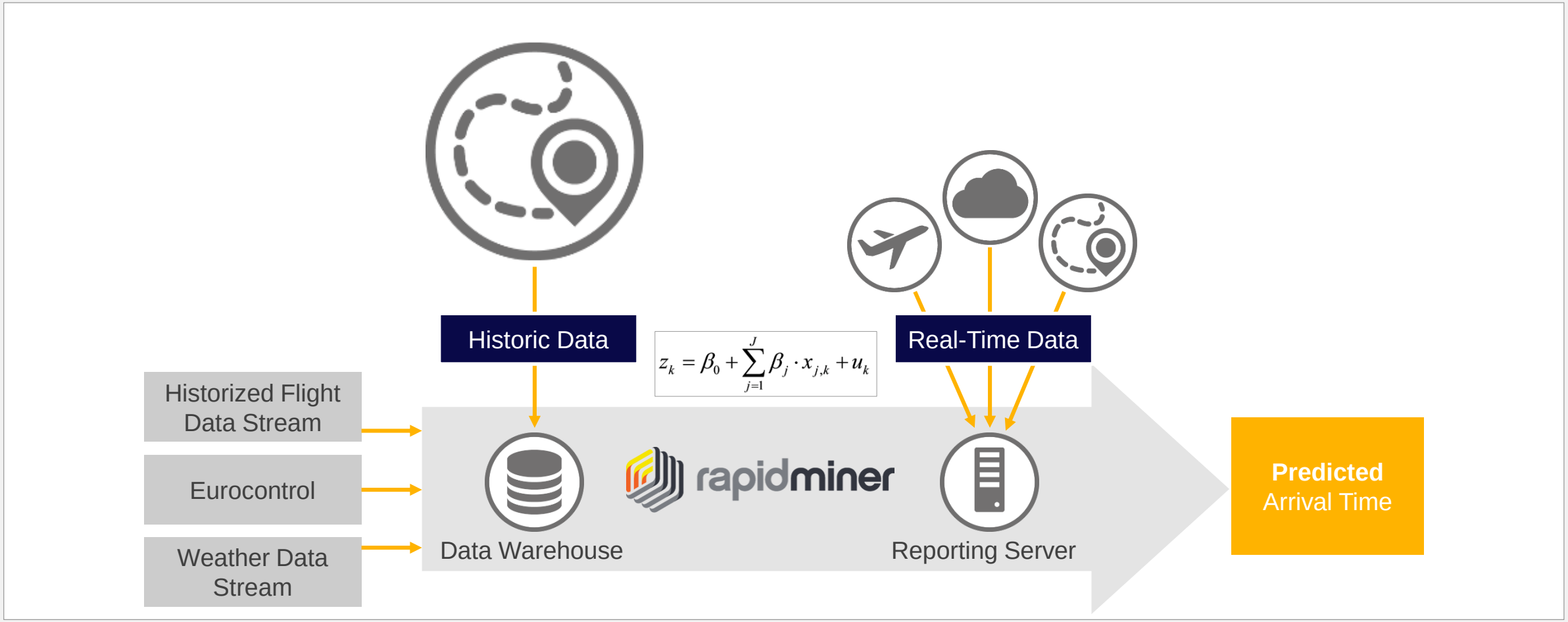
Goal & Process



Accuracy of existing prediction

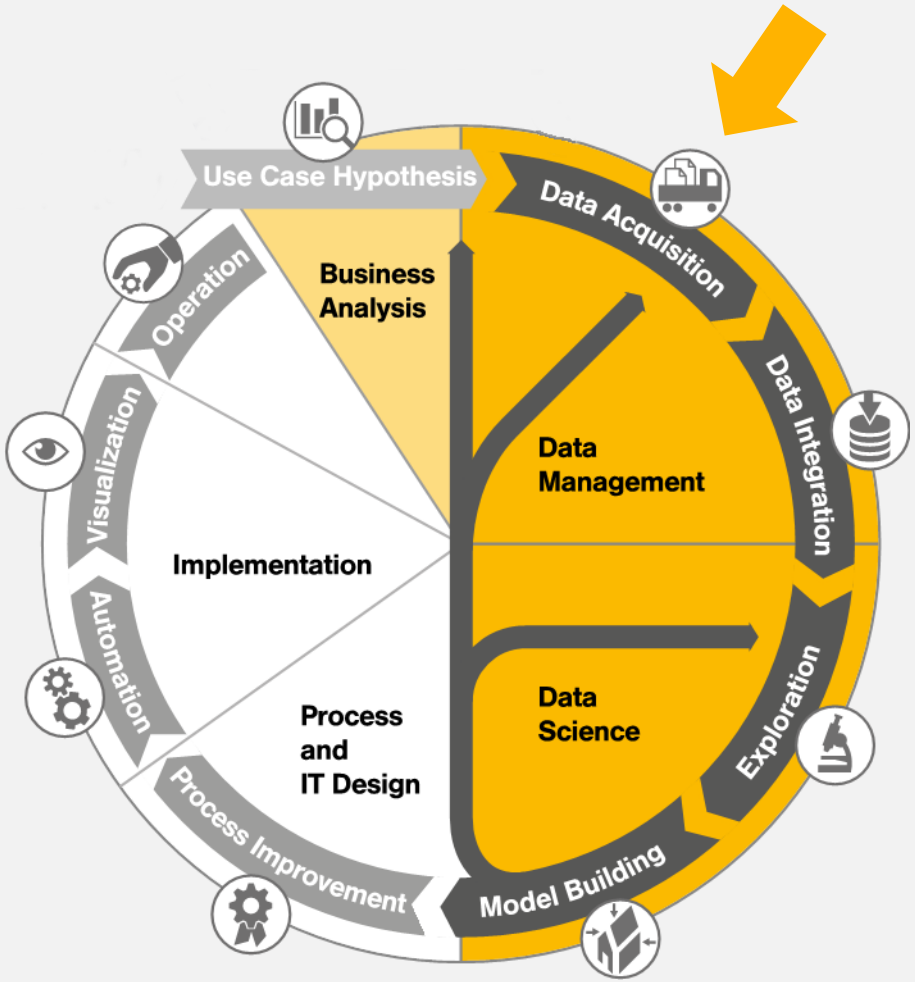


How? Unleashing The Power of Big Data



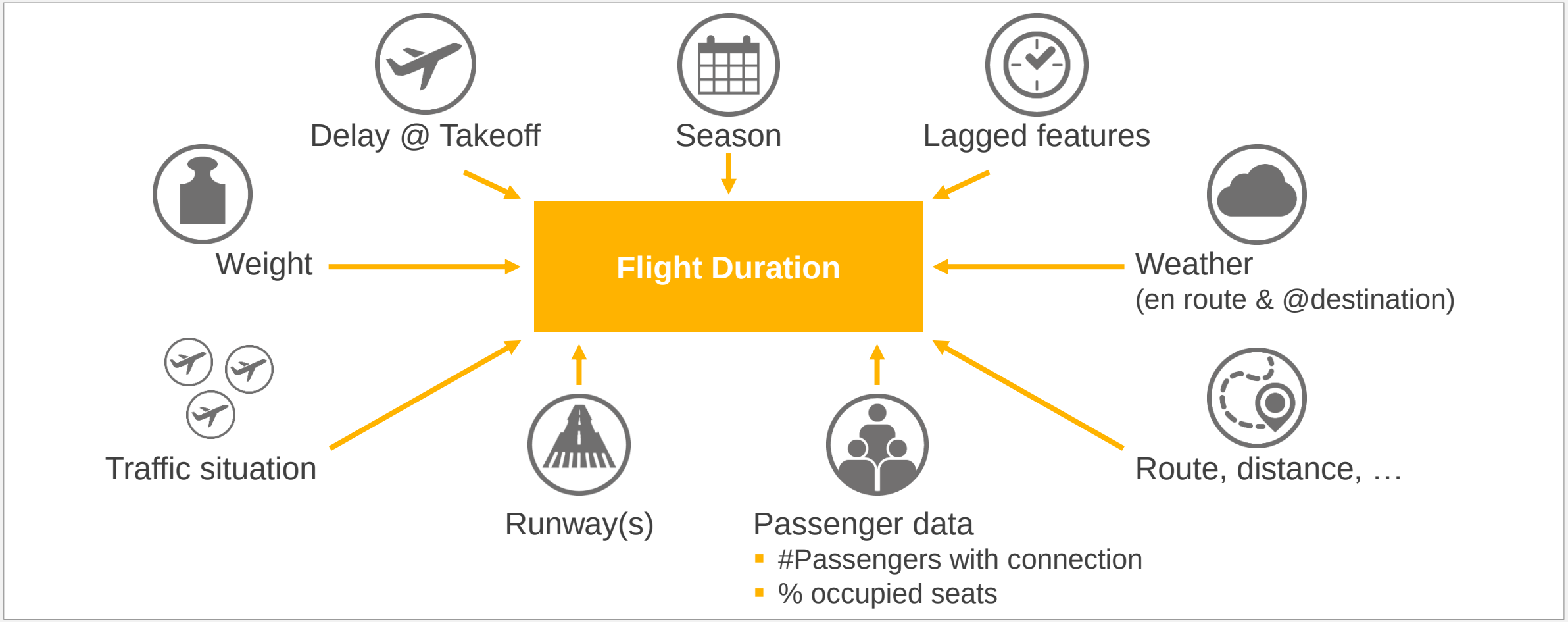
Lufthansa Industry Solutions

The Cycle

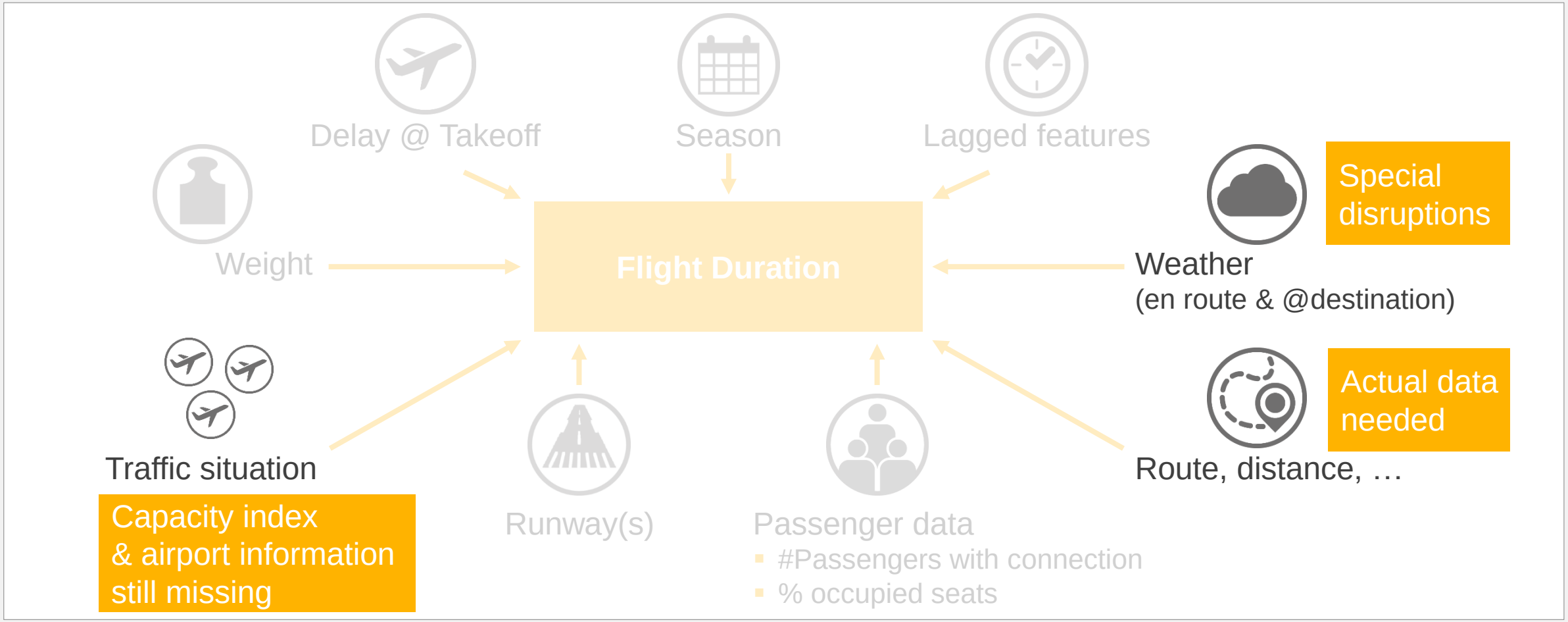


Lufthansa Industry Solutions

Parameters for the Model

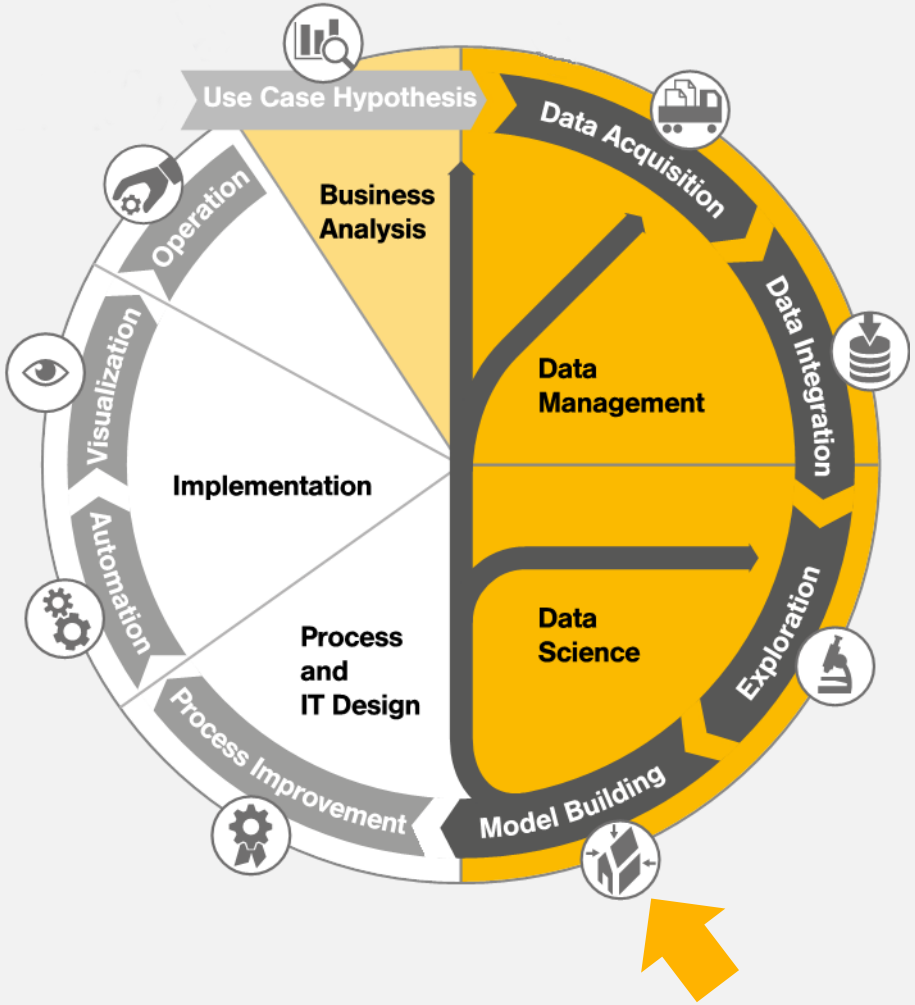


Parameters for the Model



Lufthansa Industry Solutions

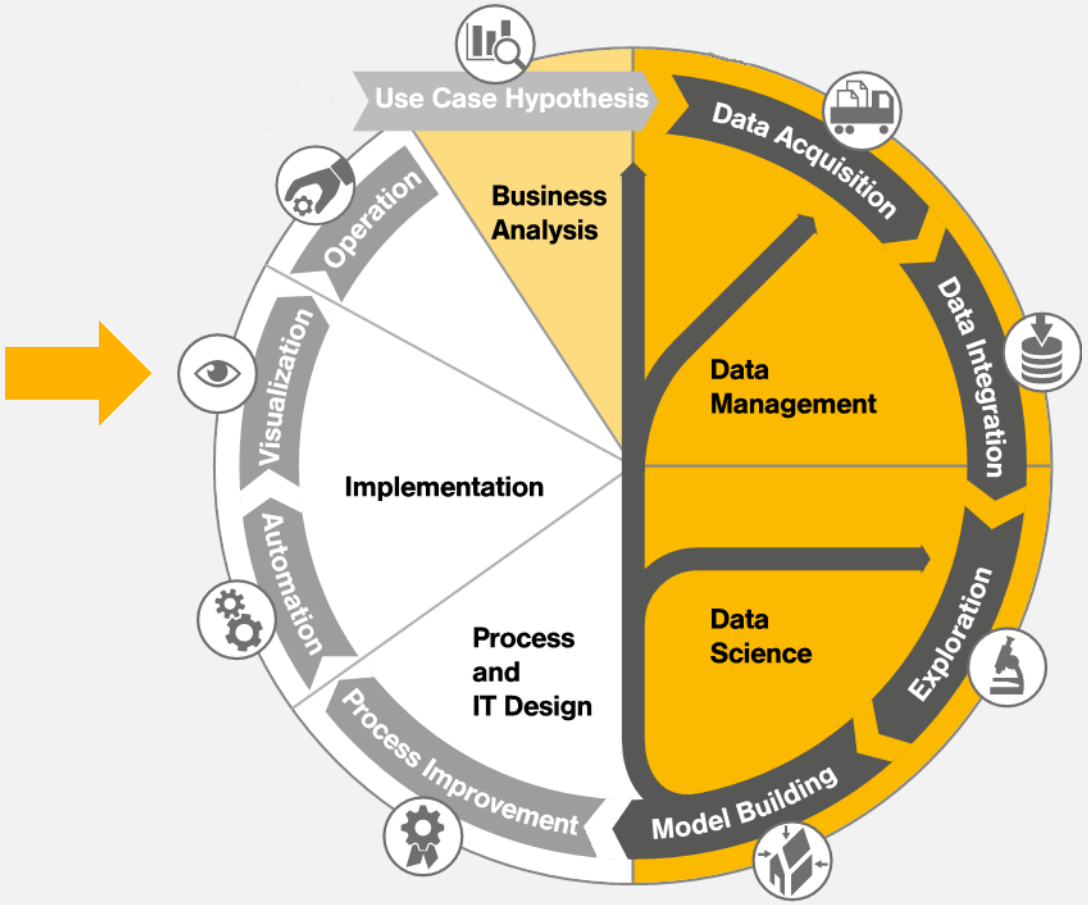
The Cycle


Use Case Hypothesis
Business Analysis
Data Acquisition
Data Integration
Data Management
Exploration
Data Science
Model Building
Process Improvement
Process and IT Design
Automation
Implementation
Visualization
Operation

Model Building

- A few rolling window features are good!
- Prefer models that can be trained fast:
  Play around with features instead of models.

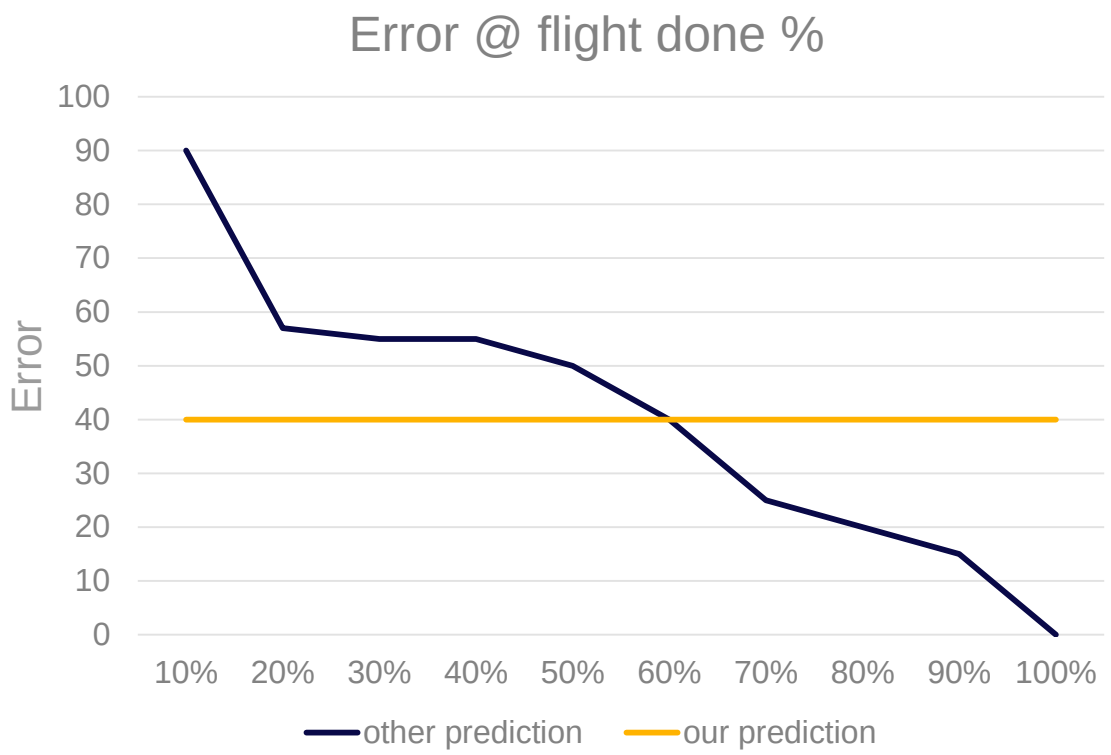
The Cycle

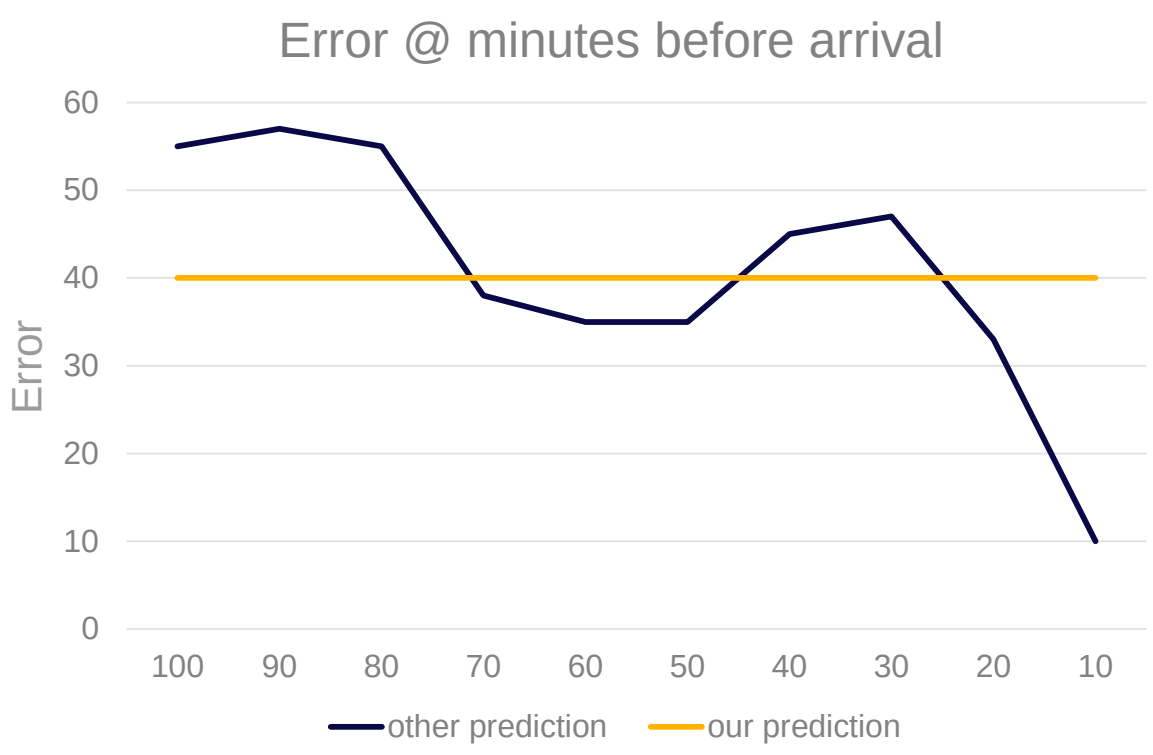


Lufthansa Industry Solutions

Visualization: Performance Monitor

Data Made Up (No Actual Data)!

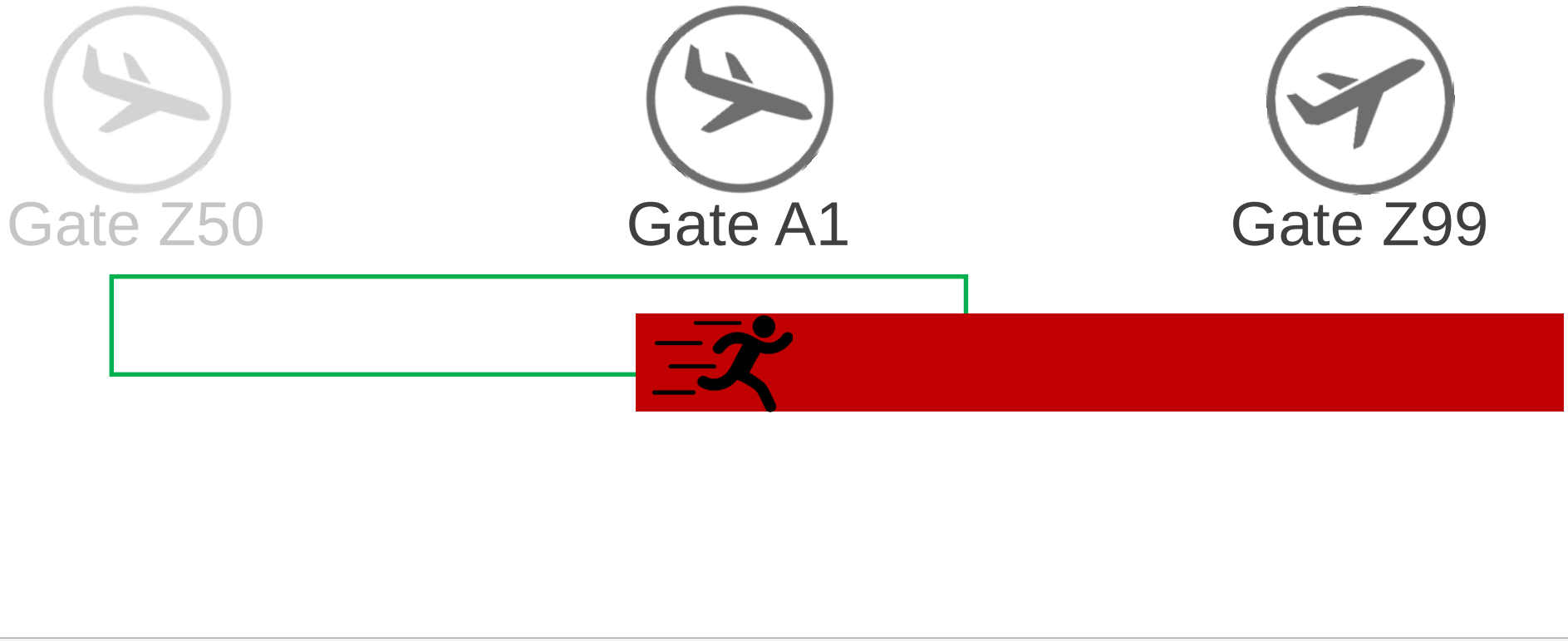




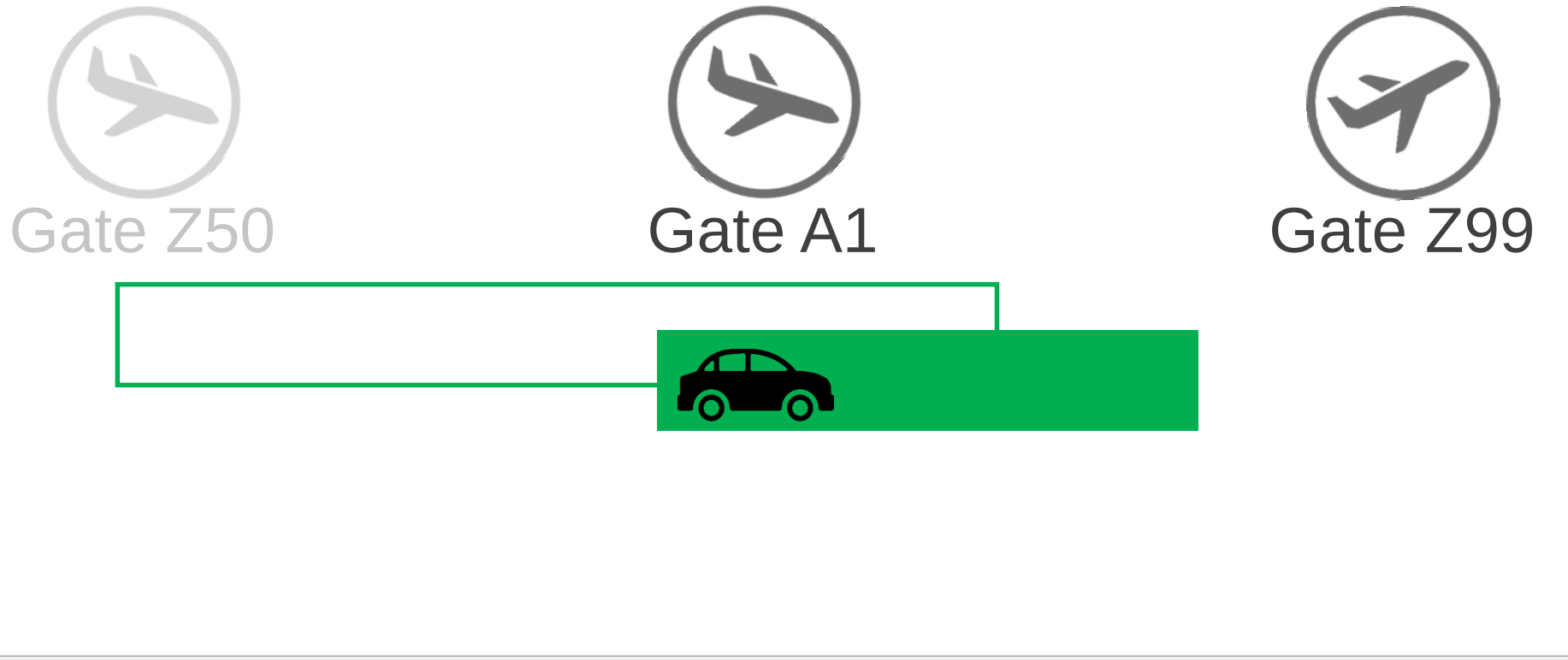
Lufthansa Industry Solutions

How To Save Money?

How To Save Money?

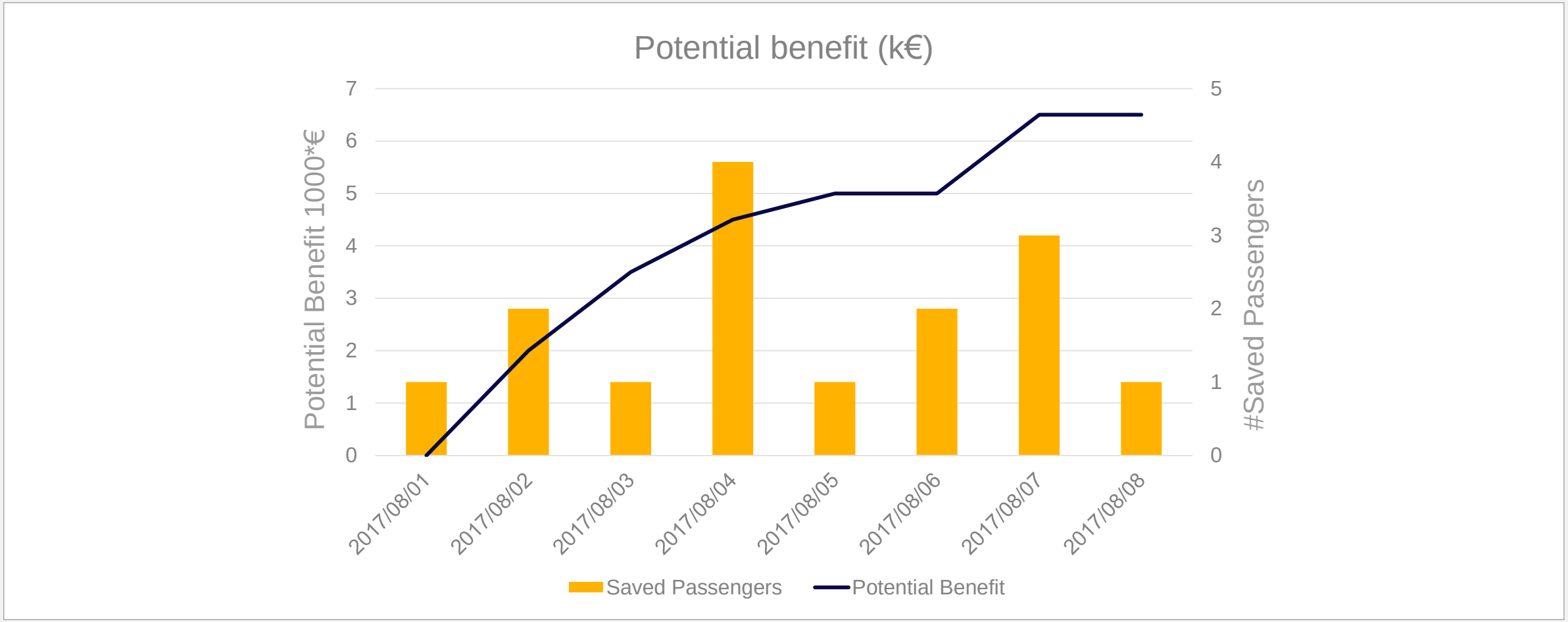


How To Save Money?



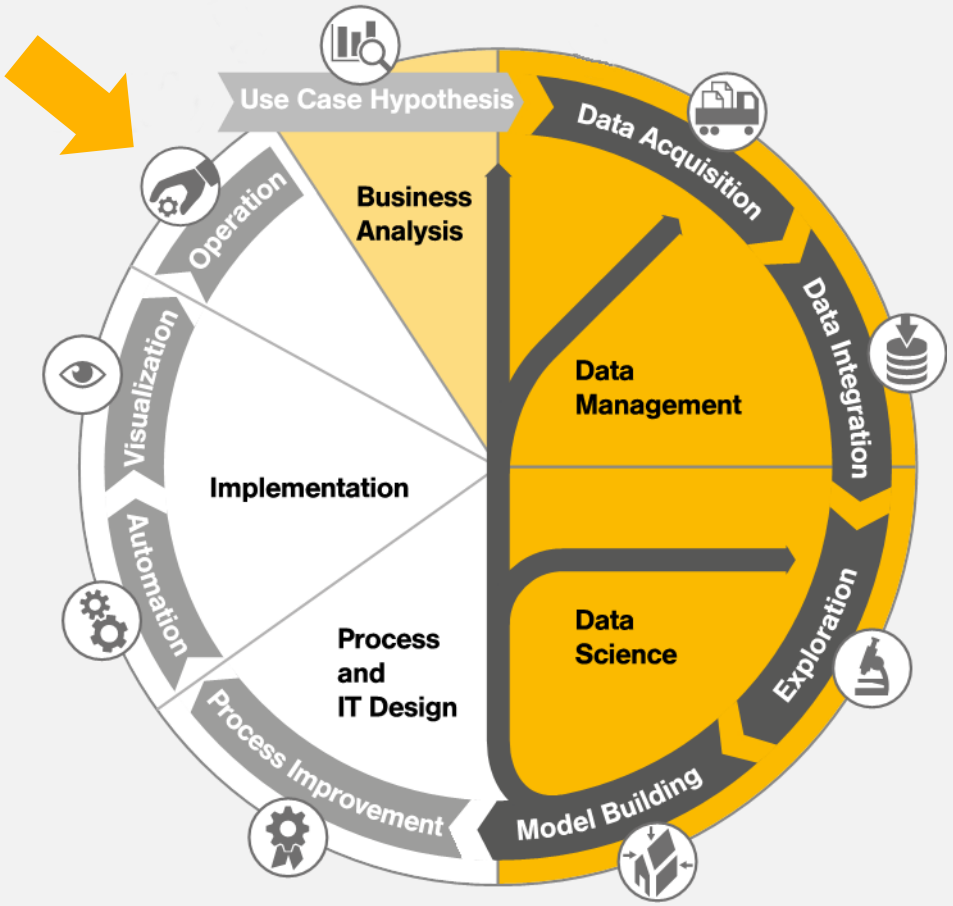
Visualization: Benefits Monitor

Data Made Up (No Actual Data)!



Lufthansa Industry Solutions

The Cycle


Use Case Hypothesis
Business Analysis
Data Acquisition
Data Integration
Data Management
Exploration
Data Science
Model Building
Process Improvement
Process and IT Design
Automation
Implementation
Visualization
Operation

Automation: Supervisor

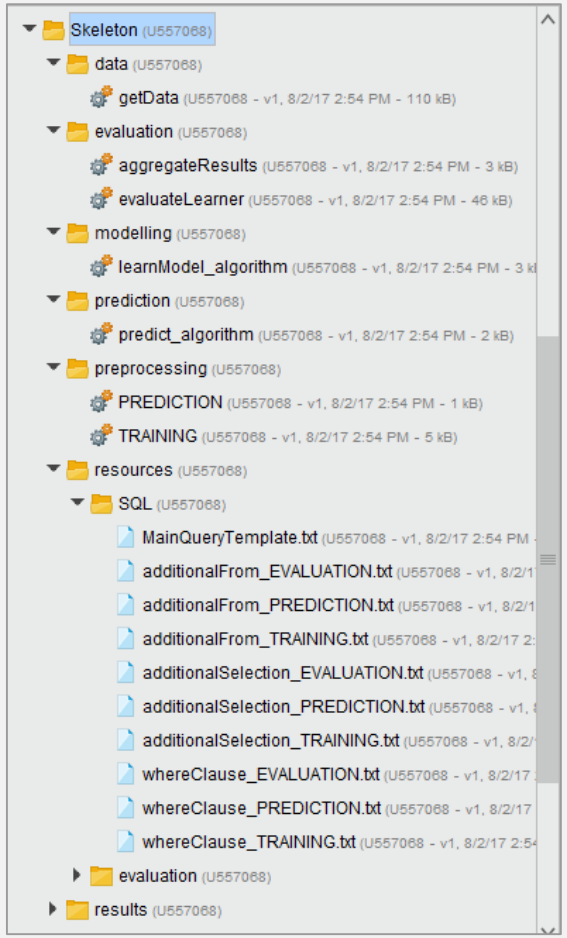
Supervision of Model

- Runs every hour
- Asserts that …
  - enough data
  - enough predictions
  - predictions produced fast
  - not too much missing data
  - …

Lessons Learned

Go To The Real World Fast – Make Everybody Understand The Benefits

- Supervisor & dashboard as early as possible
  - Business needs to understand the real impact
  - Translate Data Science to business language
- Need for standards
  - Prototype needs to be translated to real application
  - You do not want to rebuild everything



Partnership

Lufthansa Industry Solutions | RapidMiner

- Partnership founded in 2015
- RM widely used in Lufthansa
- Keeps up with market technologies
- For prototyping and productive applications
- Still lacks some software engineering capabilities …
- We pair it up with Hadoop, Oracle, Teradata, R, Python, …
- Lufthansa Industry Solutions: certified RapidMiner trainers

Lufthansa Industry Solutions

rapidminer

Lufthansa
Industry Solutions