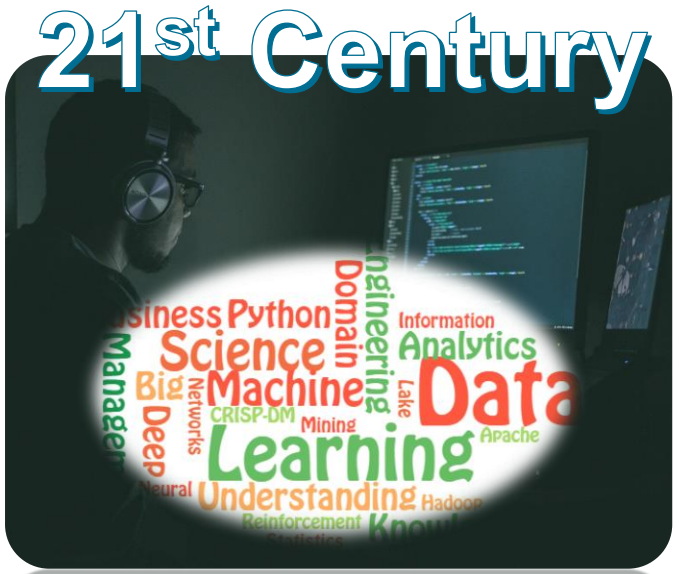
KI-Services im wissenschaftlichen und industriellen Diskurs | HWR Berlin

Engineering of Industrial AI Solutions

Dr. Jens Heidrich, Division Manager “Smart Digital Solutions”
jens.heidrich@iese.fraunhofer.de

Fraunhofer IESE

Data Scientist – “the sexiest job of the 21st century” (Harvard Business Review)



[T. Davenport and D.J.Patil: Data scientist: the sexiest job of the 21st century, Harvard Business Review, 2012]
[Pictures from https://unsplash.com, Wordle from http://www.edwordle.net]

NIST SP 1500-1, 2015

https://gi.de/datascience

Dependable AI

Fraunhofer IESE

Outline

- Why shall a company deal with AI?
- What makes AI systems special and difficult to build?
- How to engineer AI systems in industry?
- How to identify the right use cases?
- How to assure qualities of AI systems?

Why companies deal with (dependable) AI

Operational excellence

- Increasing effectiveness and efficiency of core processes
  - Preventive and predictive maintenance
  - Finding defective parts
- ⇒ Saves costs and increases revenues

Innovation

- New innovative products and services
  - Autonomous driving, collaborative robots (Cobots)
  - Platform / data-based services
- ⇒ New business models and customer groups

Customer intimacy

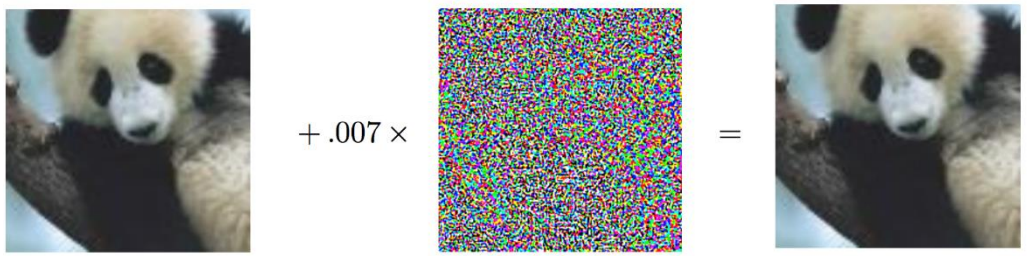
- Better understanding customers
  - Buying habits and interests
  - Custom-tailored products and offerings
- ⇒ Increases sales and revenues

Competition

- Fear of being driven out of business by companies using AI

Why is it important to talk about quality of AI systems?

Security



[Source: http://www.cleverhans.io/security/privacy/ml/2016/12/16/breaking-things-is-easy.html]

SL45 (0.77)	LE (0.04)
SL45 (0.71)	STP (0.08)
SL45 (0.47)	STP (0.30)
SL45 (0.79)	STP (0.05)
SL45 (0.79)	STP (0.06)
SL45 (0.68)	STP (0.12)
SL45 (0.67)	STP (0.11)

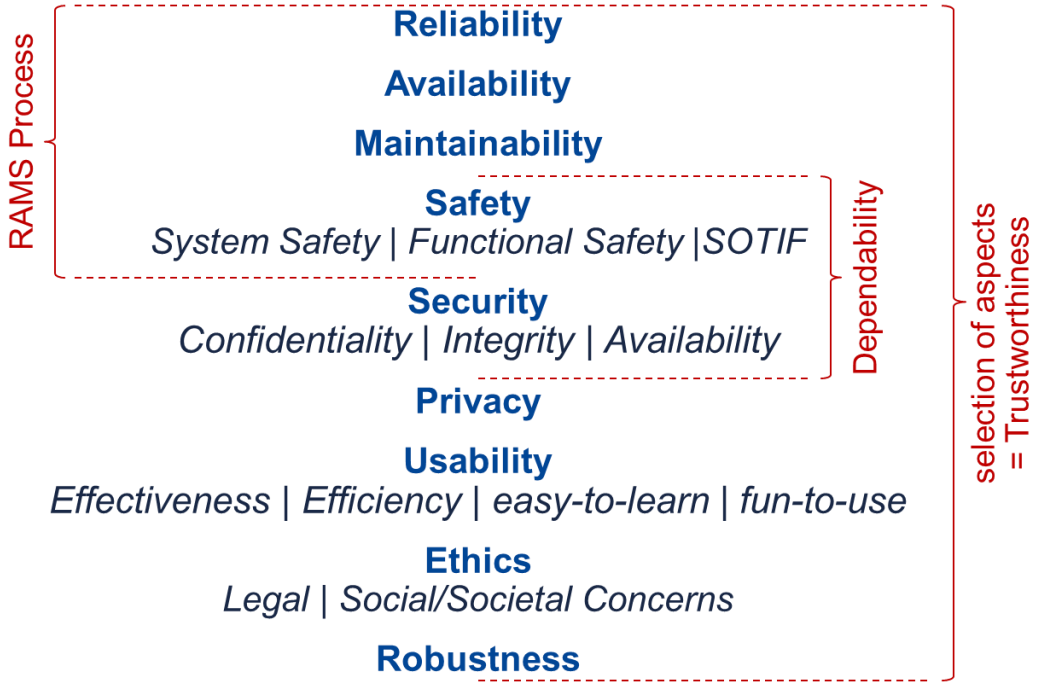
[Source: Ivan Evtimov, et al.: „Robust Physical-World Attacks on Machine Learning Models", arXiv:1707.08945v5, 04/2018]

Safety

[Source: https://www.theregister.co.uk/2017/06/20/tesla_death_crash_accident_report_ntsb/]

[Source: https://www.thenational.ae/business/uber-turned-off-volvo-crash-prevention-system-before-fatal-accident-1.716390]

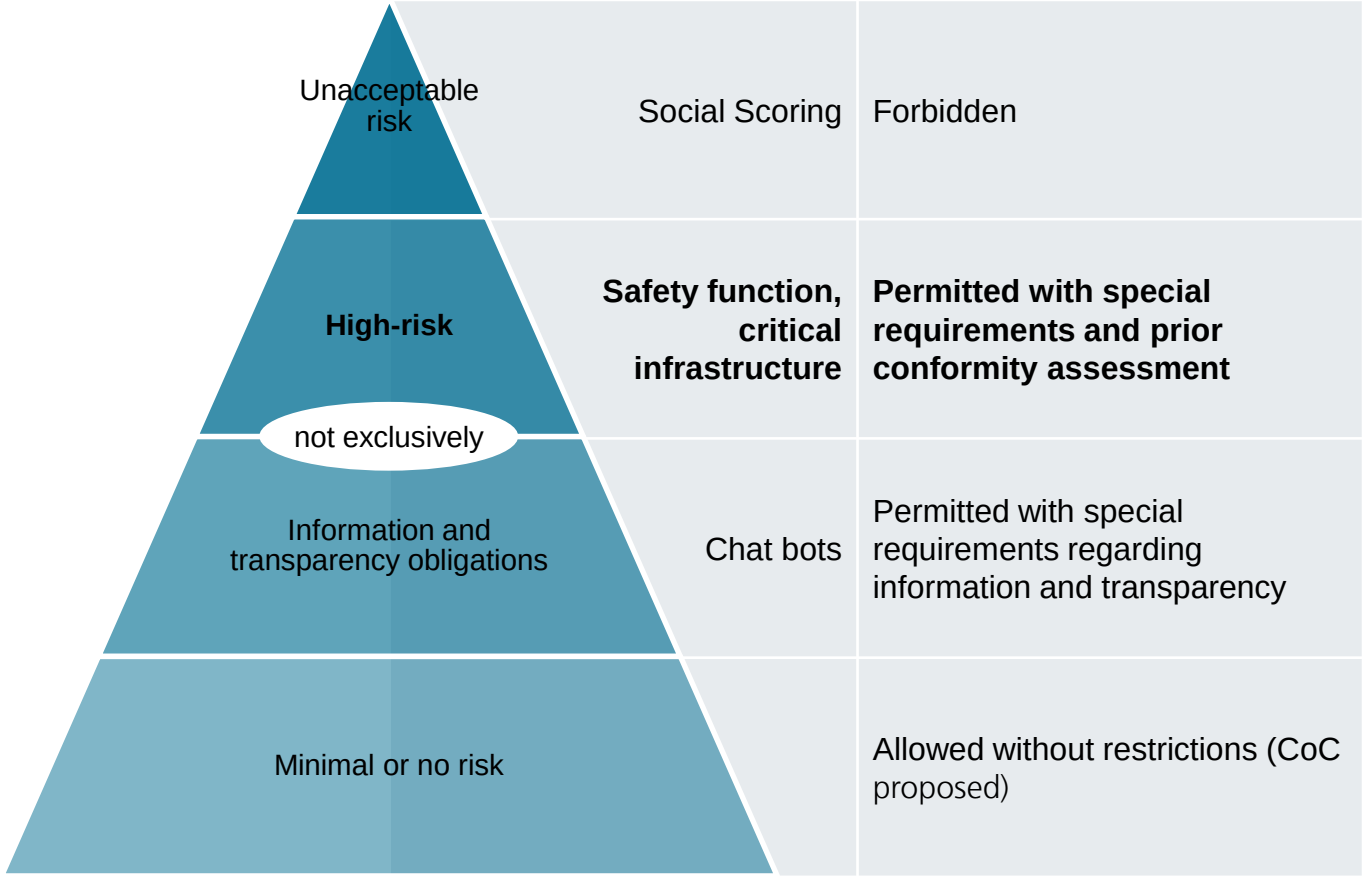
Qualities of AI Systems and specifically Dependability of AI Systems



[Source: VDE application rule VDE-AR-E 2842-61, DKE/AK 801.0.8 material for the AR "Development and Trustworthiness of autonomous/cognitive Systems"]

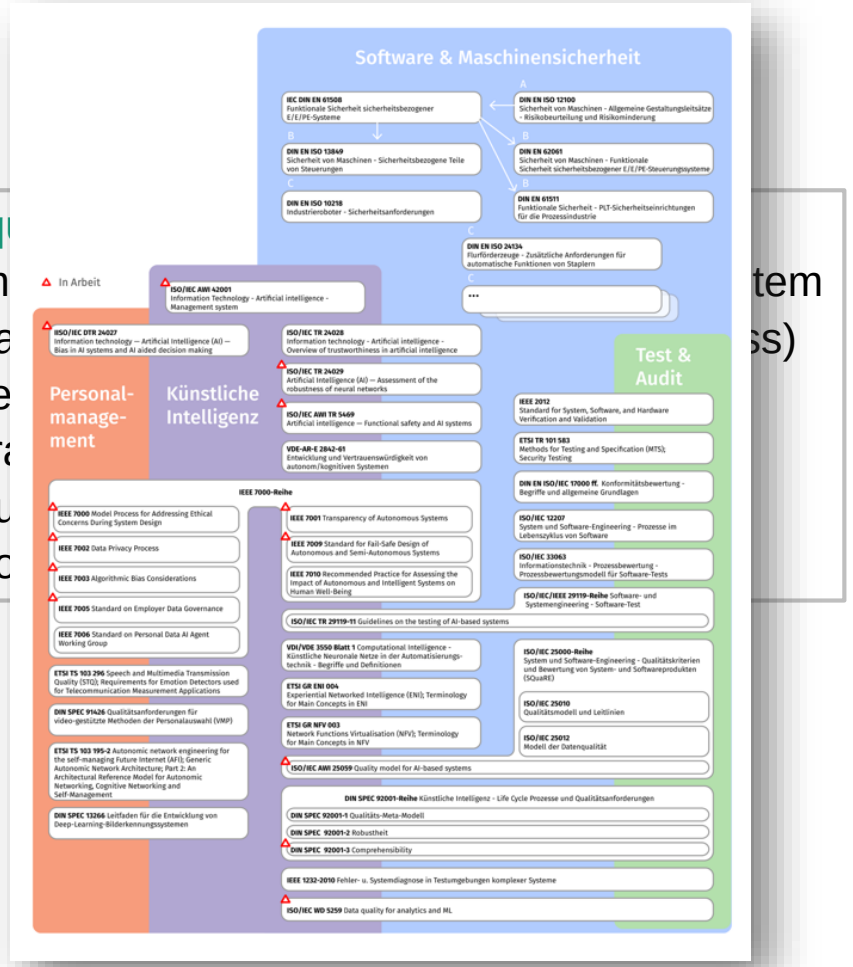
- **Dependability** of a system describes "its ability to avoid unacceptable failures in the provision of a service or functionality" (Jean-Claude Laprie)
- Dependability is crucial when it comes to using AI in the context of critical application fields, such as:
  - **Mobility and logistics:** autonomous driving functions, traffic management systems, etc.
  - **Industrie 4.0:** collaborative robots, driverless transport systems, etc.
  - **Digital Health:** prevention, diagnosis, and therapy of diseases, surgery robots, etc.
  - **Smart Energy:** energy controlling and management systems, service robots, etc.
- For those kind of AI systems the use of AI may bear a **high risk** for direct / indirect personal casualty

EU AI Act: Classification of AI Systems and Requirements based on Risk



	Examples	Regulation
Unacceptable risk	Social Scoring	Forbidden
High-risk	**Safety function, critical infrastructure**	**Permitted with special requirements and prior conformity assessment**
not exclusively		
Information and transparency obligations	Chat bots	Permitted with special requirements regarding information and transparency
Minimal or no risk		Allowed without restrictions (CoC proposed)

Example: Regulations and Standards in Production Domain



Dependable AI

Fraunhofer IESE

Typical Industry Challenges for Engineering AI Systems

AI System Engineering Process

Ramp-Up and Ideation

- Limited data science and software engineering competencies
- Finding the right use case and business case
- Availability of data

Construction and V&V

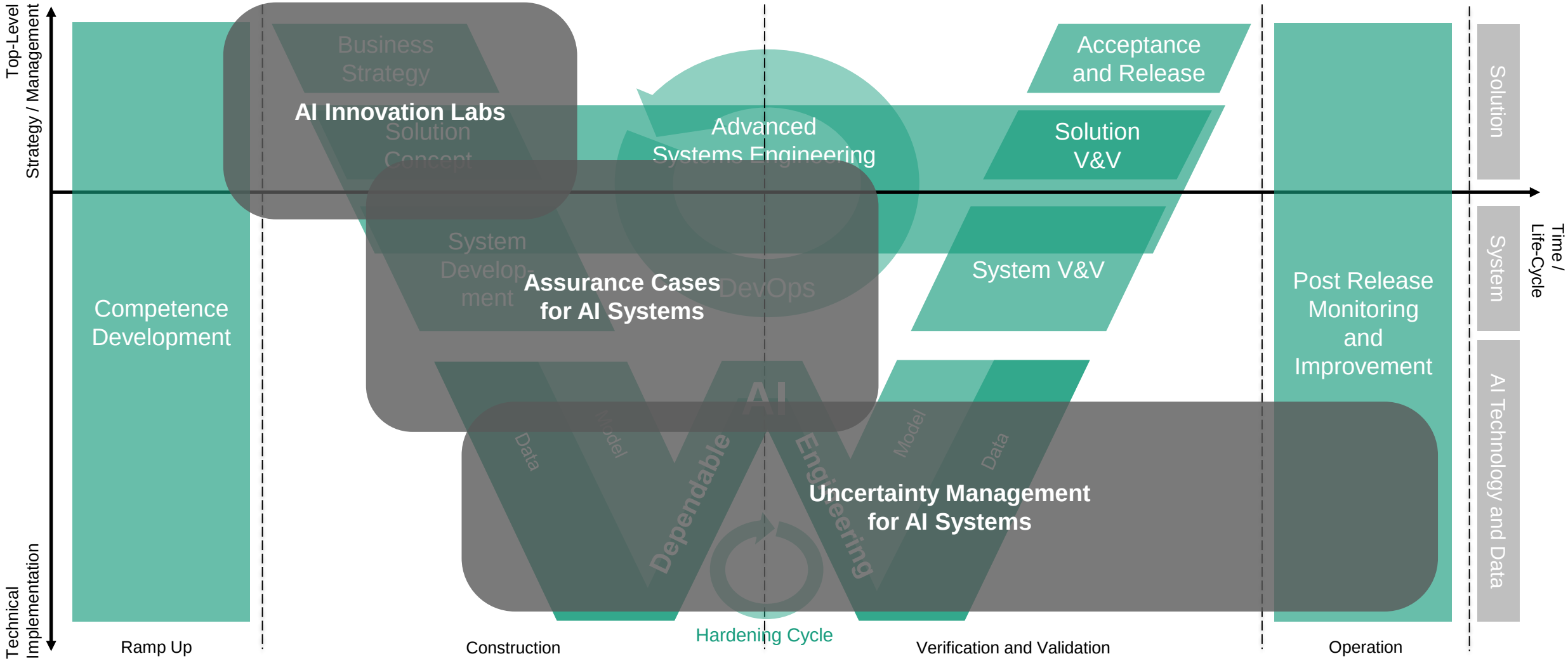
- Build product from prototype
- Unclear how to proof compliance to regulations and standards
- Approaches for testing and certification
- Easy-to-use development environments for AI systems

Operation

- Observation and management of AI performance at runtime
- Maintenance of AI models

[Partially based on S. Wrobel, Fraunhofer IAIS, Fraunhofer Technologietag, Stuttgart, February 2019]

Overview of Process for the Engineering of AI Systems

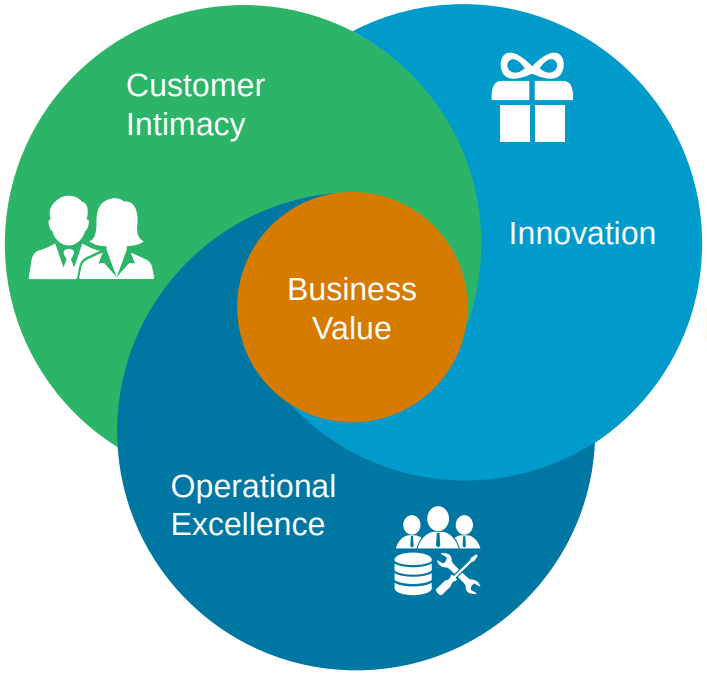


[VDE-AR-E 2842-61: A reference model for trustworthy AI, © DKE]

Dependable AI

Fraunhofer IESE

AI Innovation Labs: Business Solution



Key Partners
- Injection molding production companies
- Machine manufacturer

Key A
- Build predictive maintenance software
- Assure safety

Key Resources
- Machine Data
- Data Scientists

Value Propositions
- 24/7 production
- Maintain production machine before critical failure
- Avoid production stops

Customer Relationships
- Co-creation

Channels
- Direct channel

Customer Segments
- Injection molding production

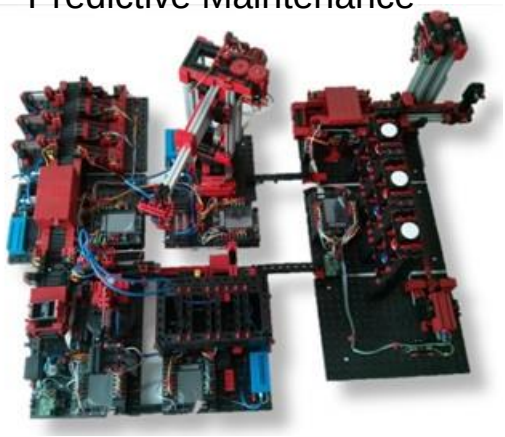
Cost Structure
- Additional HW (Sensors)
- Software Development Cost

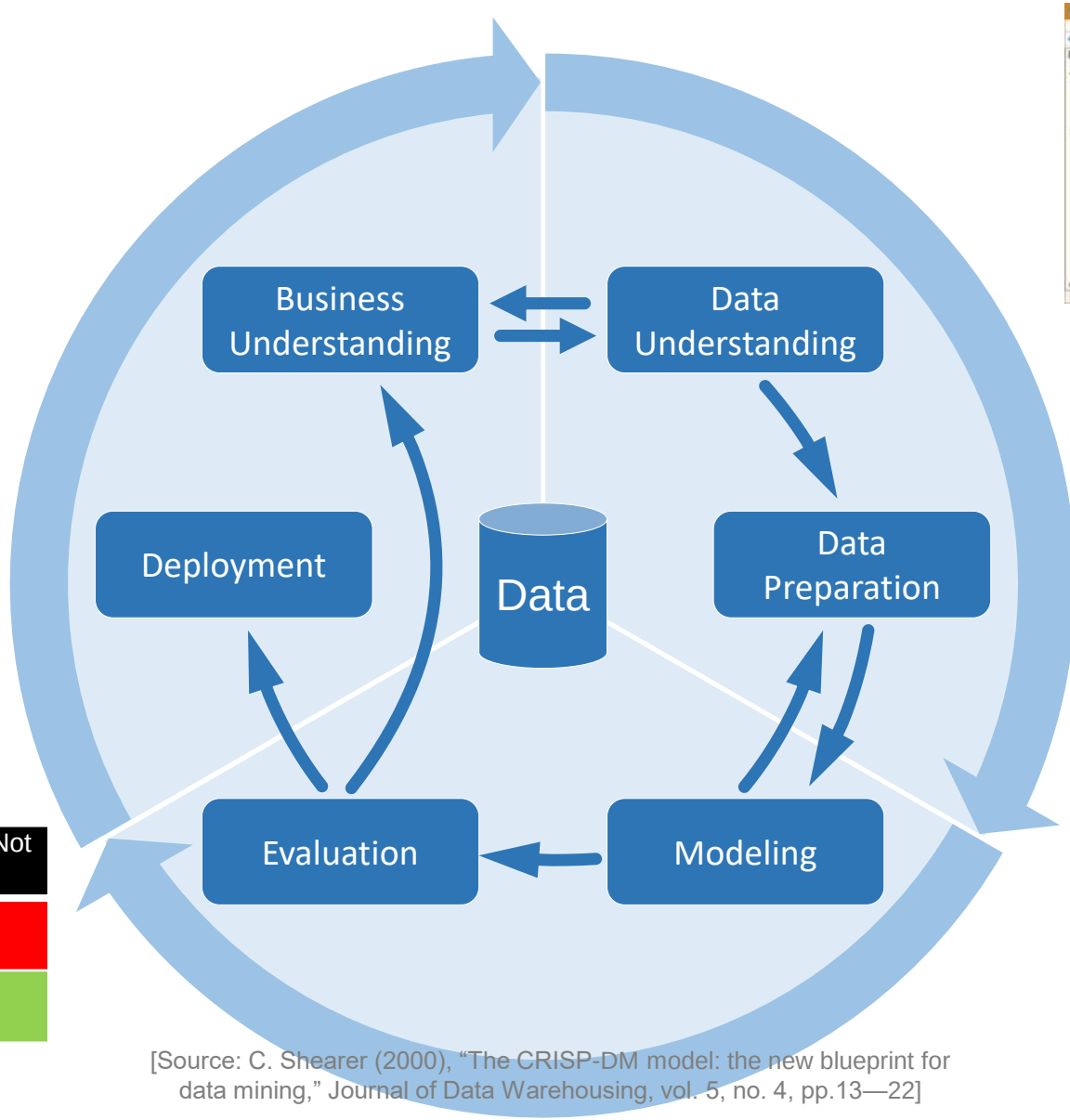
Revenue Streams
- X% more uptime of machines ⇒ Y% more profits

[Source: © Business Model Canvas by strategizer.com]

AI Innovation Labs: Technical Solution

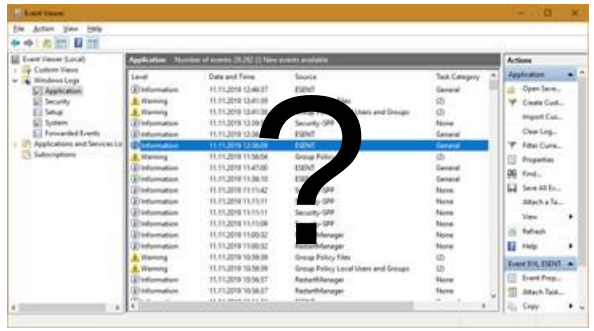
Predictive Maintenance





[Source: C. Shearer (2000), "The CRISP-DM model: the new blueprint for data mining," Journal of Data Warehousing, vol. 5, no. 4, pp.13—22]

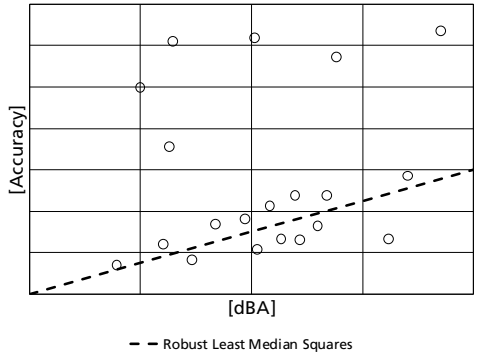
Machine Log File



Training and Test Data

Time	Angle	Noise	Accuracy
2019-11-11 06:22:23	5°	100 dBA	99%
2019-11-11 06:22:33	8°	99 dBA	99%
2019-11-11 06:22:43	30°	110 dBA	99%
2019-11-11 06:22:53	45°	200 dBA	70%
2019-11-11 06:23:03	8°	101 dBA	99%
...			

Simple Regression Model



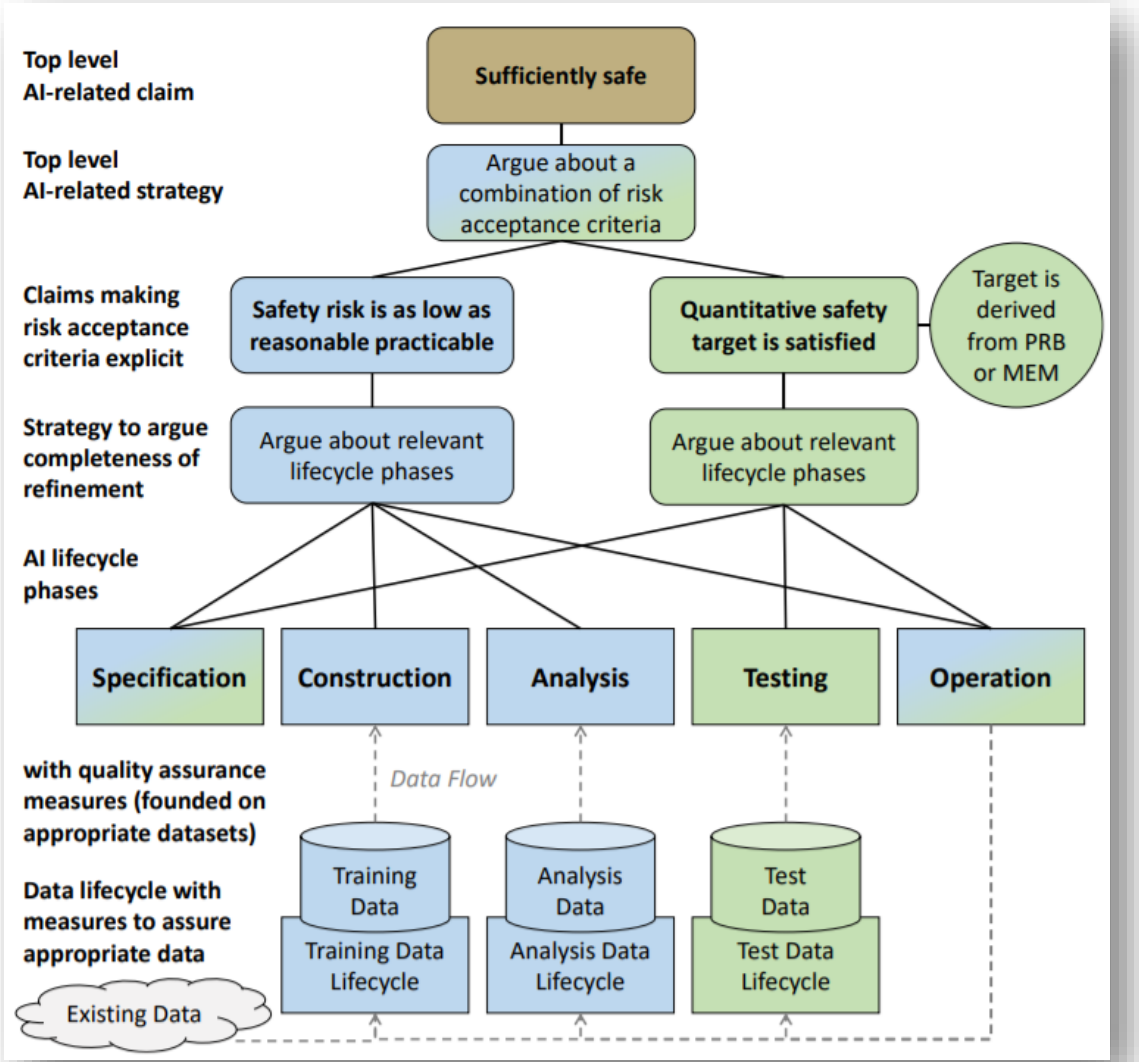
Model Accuracy

Truth / Predicted	Positive (Defective)	Negative (Not Defective)
Positive (Defective)	True Positives	False Positives
Negative (Not Defective)	False Negatives	True Negatives

Assurance Cases for AI Systems

- Upcoming regulations demand certification of high-risk AI
- Currently, no standards exist that can easily be applied
- **Assurance cases** are a structured chain of arguments with associated evidence that allows the assumption that a product in a certain usage environment meets the set goals (such as safety)
- They are well-known concepts from the **safety engineering** domain and have proven to be applicable for AI systems in different domains
- Currently, they are seen as the best practice for arguing about the dependability of an AI System!

[Source: Kläs, M., et al.,"Using Complementary Risk Acceptance Criteria to Structure Assurance Cases for Safety-Critical AI Components," AI Safety 2021 at International Joint Conference on Artifical Intelligence (IJCAI), Montreal, Candada, 2021.]



Uncertainty Management for AI Components

- Uncertainty is inherent in data-based solutions and its sources must be clearly identified, quantified and managed
- **Uncertainty Wrappers** allow for identification and estimation of uncertainty in AI components
- Used for generating evidences in an assurance case
- Uncertainty wrappers at runtime allow the system to go to a safe state if the **estimated** uncertainty is too high

Outcome?
No Stop Sign

Uncertainty?
0.02 / 0.40

Confidence?
0.9999

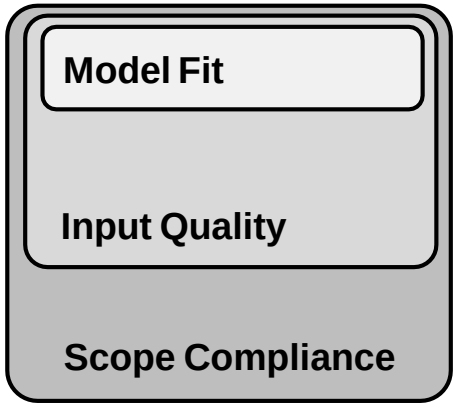
[Source: Kläs, M., et al.,"Handling Uncertainties of Data-Driven Models in Compliance with Safety Constraints for Autonomous Behaviour," Proceedings of European Dependable Computing Conference (EDCC 2021), Munich, Germany, IEEE, 2021.]

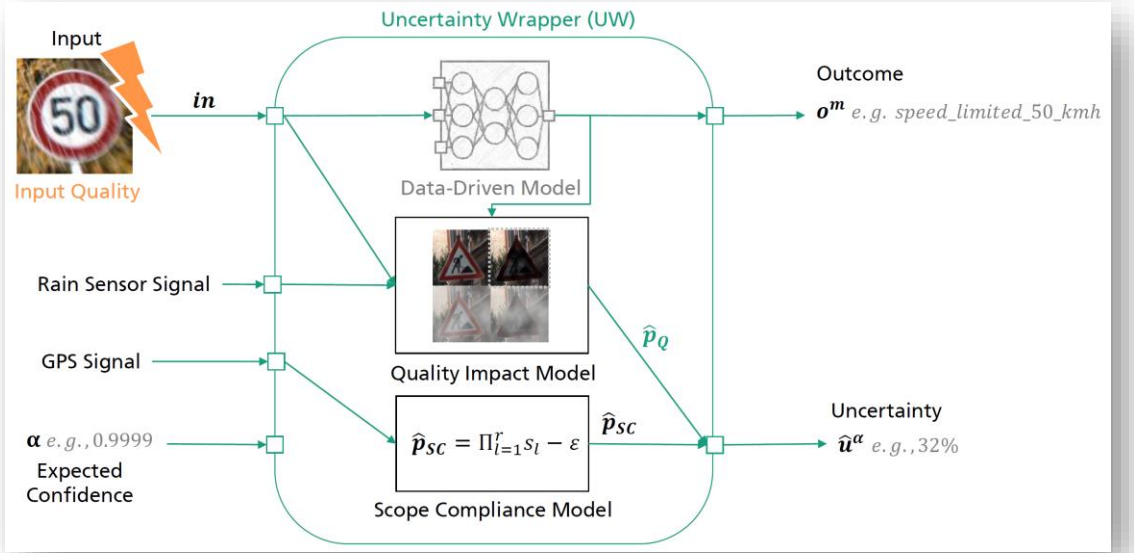
Sources of Uncertainty — *Onion Shell Layers*

(inherent) limitations of the learned model

+ input quality limitations during model application

+ mismatch with target application scope





Conclusions

- AI systems bear a lot of **potential**
- Engineering AI systems is **challenging**
  - Classic methods for V&V hardly applicable
  - Methods and tools only partially available (part of research)
- Building AI systems require a **proper engineering process**
  - Think about **use cases** and benefits before technologies and data lakes
  - Follow on iterative, **prototyping-oriented process** for trying out new ideas
  - Identify and **assure quality goals** of AI systems
  - **Master uncertainty** of AI systems

