Bundesrechnungshof

ECA conference on Big and Open Data
for European supreme audit institutions

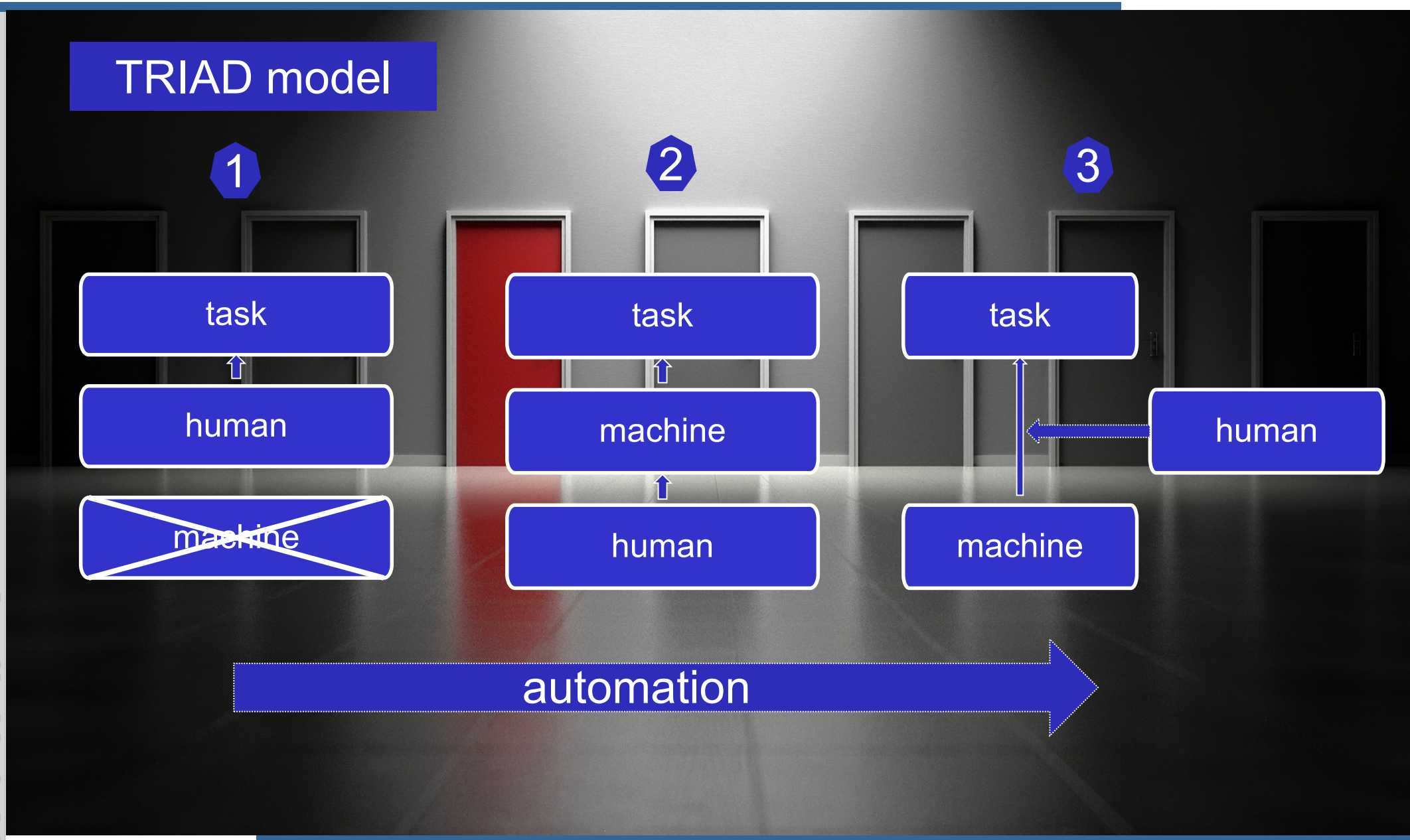
Luxembourg, 27-28 November 2019

# Auditing of automated decision-making

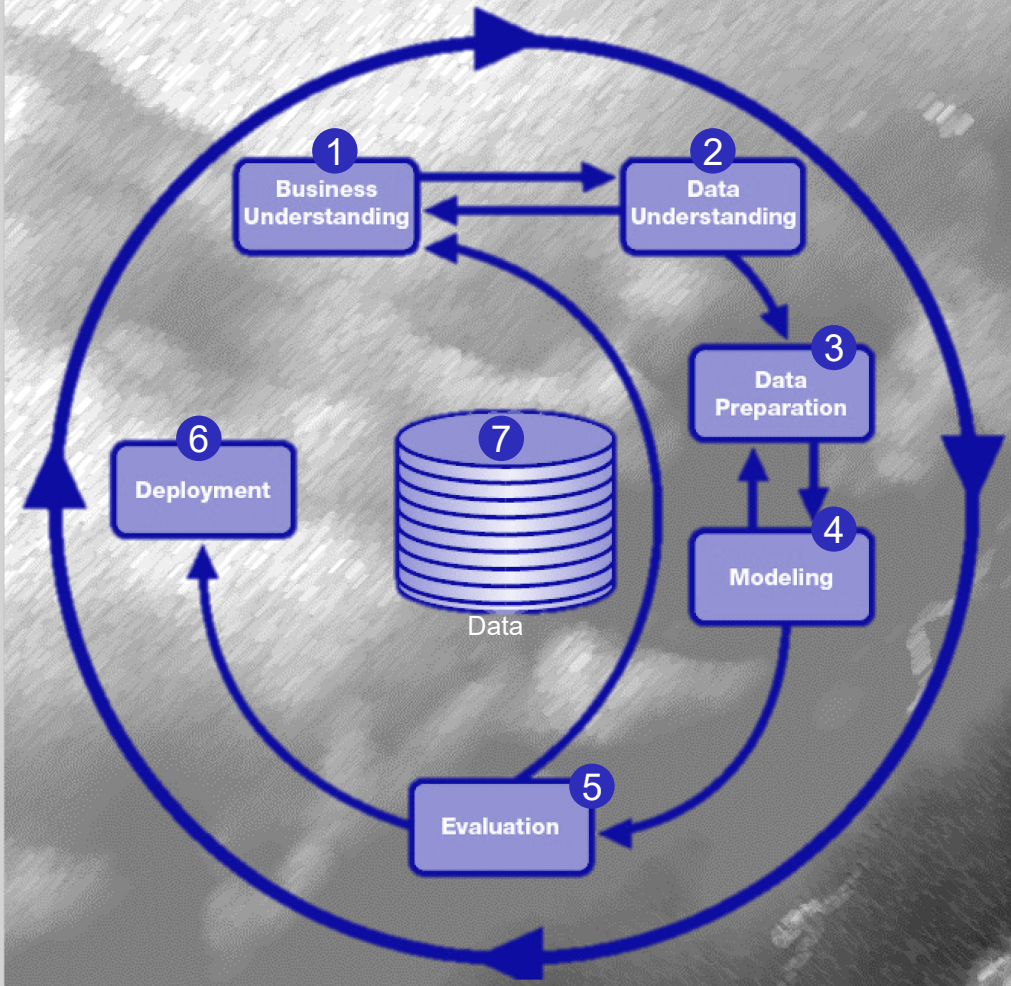
## Challenges of deep learning and data analytics from the auditor‘s point of view

(Peter Koß, audit manager, German SAI)

# Decision-making as the goal of data analytics



# CRISP-DM - a guideline for the audit process




Each phase of the data analytics process as described by CRISP-DM is well-suited for stepping in with audit work

Prepare audit goals and activities for:

- Business understanding
- Data understanding
- Data preparation
- Modeling
- Evaluation
- Deployment
- Data provision



Peter Koß, audit manager, German SAI

# CRISP-DM - phases and the auditor's capabilities

Bundesrechnungshof

Each CRISP-DM phase requires different auditing abilities and audit knowledge:

Level of complexity increases

| | |
|---|---|
| (1-2) | • be familiar with the field of business |
| (3,7) | • have an understanding of database technologies, data and quality management |
| (4-5) | • have a track record in data analytics, statistics, and quality measures |
| (6) | • be familiar with software architectures, deployment, and data procurement |

# A look into the black box

- Machine learning models are “black boxes”
  - Coherencies learnt are complex and abstract
  - Models depend on data used
  - Model evaluation
    - Statistical evaluation
    - Semantical evaluation
  - Model performance is evaluated by means of standard metrics
  - Metrics may lead to erroneous conclusions about model quality
- Humans do not understand the decision-making process

# A look into the black box - model evaluation

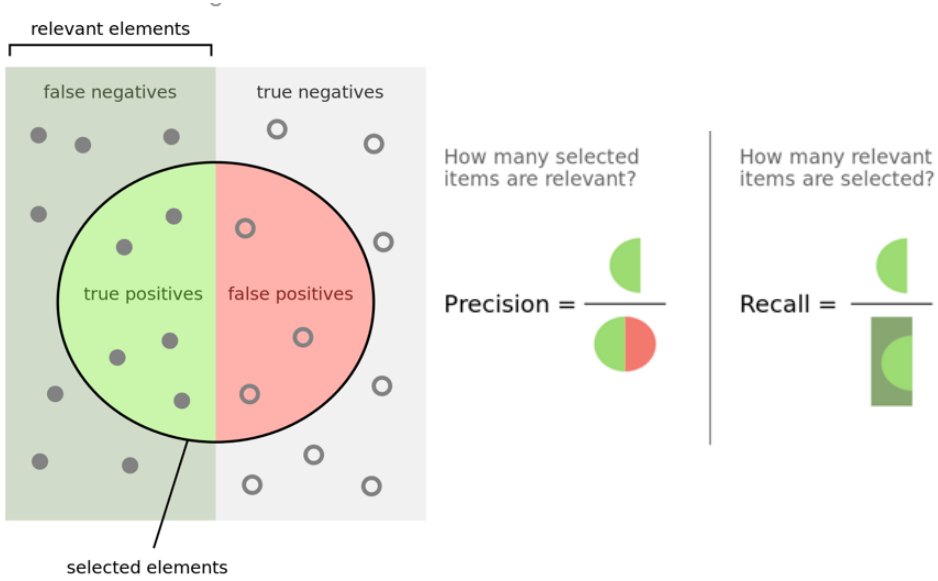
Bundesrechnungshof

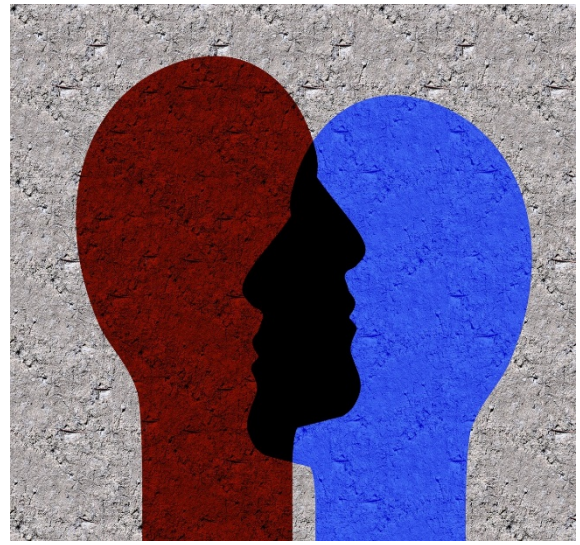
## Statistical evaluation

- Evaluate blind tests
- Evaluate statistical tests
  - Classification
  - Regression
  - Clustering

## Semantical evaluation

- Understanding of the results
- Alignment with business goals
- Plausibility of assumed correlations
- Implications of bias, variance and random noise
- Means for prevention of overfitting



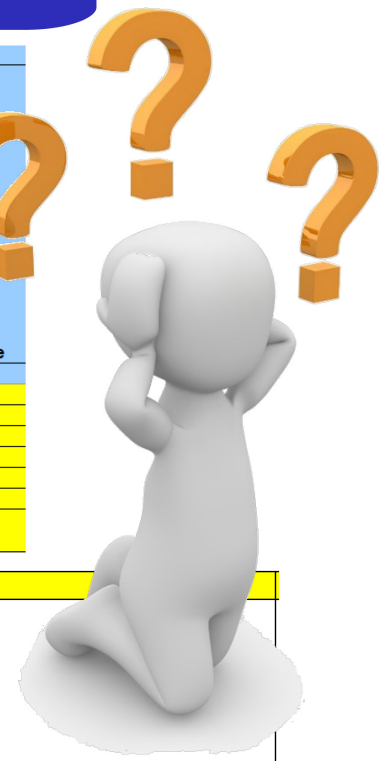
# Audit catalogue for data mining and AI systems

Bundesrechnungshof

Audit catalogue derived from CRISP-DM containing a total of 71 audit questions

| No. | Relevance | Priority | Type | Questions based on the process model: Cross Industry Standard Process Data Mining (CRISP-DM) | Purpose of the question | Document needed as audit evidence (if appropriate and available) | Sys Roles concerned: Autonomy and control | Sys Roles concerned: Requisitioning unit | Done (yes="X", no=empty case) | Response |
|---|---|---|---|---|---|---|---|---|---|---|
| 71 | | | | | | | | | | |
| A1 | | | | Microeconomic questions | | | x | x | 0% | |
| A2 | | | | Understanding of data | | | | x | 0% | |
| A3 | | | | Data preparation | | | x | x | 0% | |
| A4 | | | | Modelling | | | x | x | 0% | |
| A5 | | | | Evaluation | | | x | x | 0% | |
| A6 | | | | Deployment and operation | | | x | x | 0% | |
| A7 | | | | Management processes during development, rollout and operation | | | | x | 0% | |

| | | | | | | |
|---|---|---|---|---|---|---|
| A5 | | | | Evaluation | | |
| A5.001 | | | | Specify the validation methods you applied and state the underlying criteria. | - Identify,<br>--- how the model quality has been checked,<br>--- how the system decisions/forecasts have been tracked,<br>--- how the influence of individual criteria on the decisions has been analysed,<br>--- how it has been checked whether criteria not taken into account might have led to better decisions, and<br>--- how the fairness of the model has been checked. | - description of validation methodology used |



For audit work on development, rollout, and operation of AI and data mining systems

We are currently piloting the catalogue in close cooperation with the Federal Employment Agency



Peter Koß, audit manager, German SAI

# Key takeaways

Bundesrechnungshof

- Automated decision-making presents new challenges for our audit work
- The CRISP-DM standard offers audit intervention points at each phase
- Auditors need specific skills and tools to audit automated decision-making
- Statistical or semantical evaluation may be applied to “black box” models
- Our new audit catalogue covers a broad range of audit questions and topics

**Thank you for your attention.**



Peter Koß, audit manager, German SAI