LET'S TALK ABOUT PALM LEAVES
FROM MINIMAL DATA
TO TEXT UNDERSTANDING

MAGNUS BENDER[1], MARCEL GEHRKE[1], TANYA BRAUN[2]
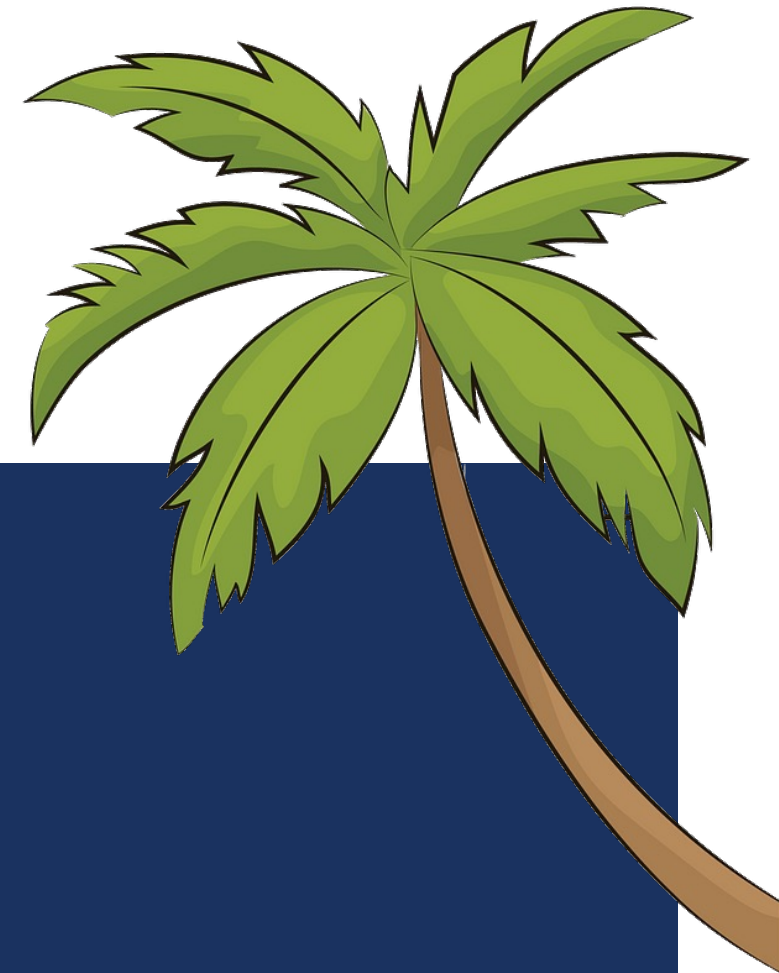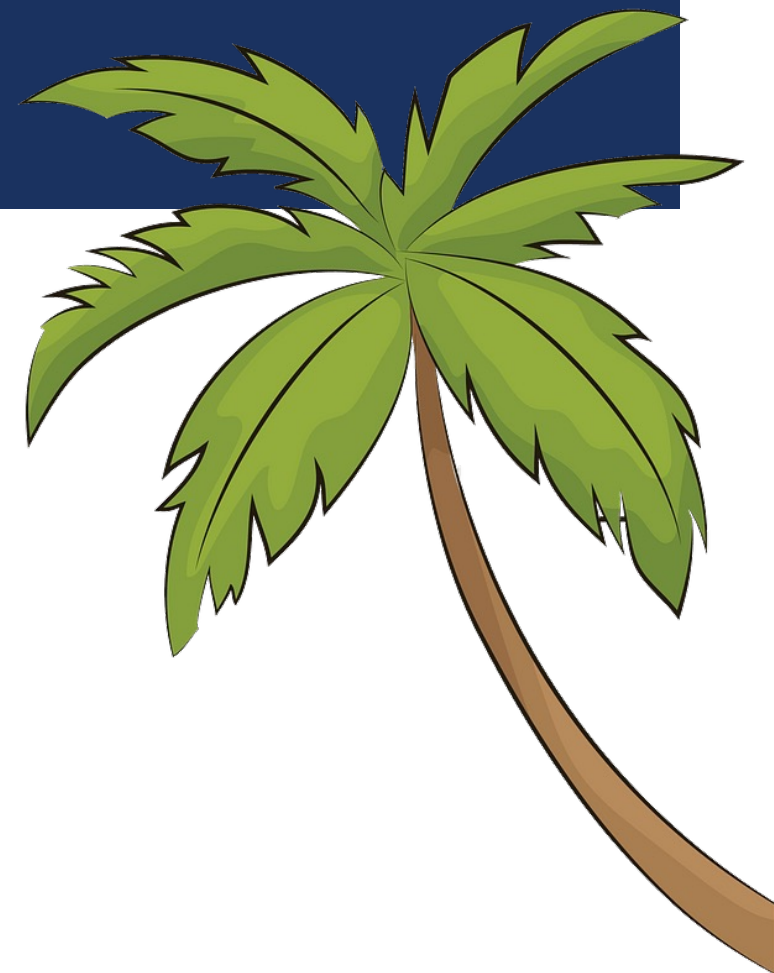
[1]Institute of Information Systems, University of Lübeck
[2]Computer Science Department, University of Münster
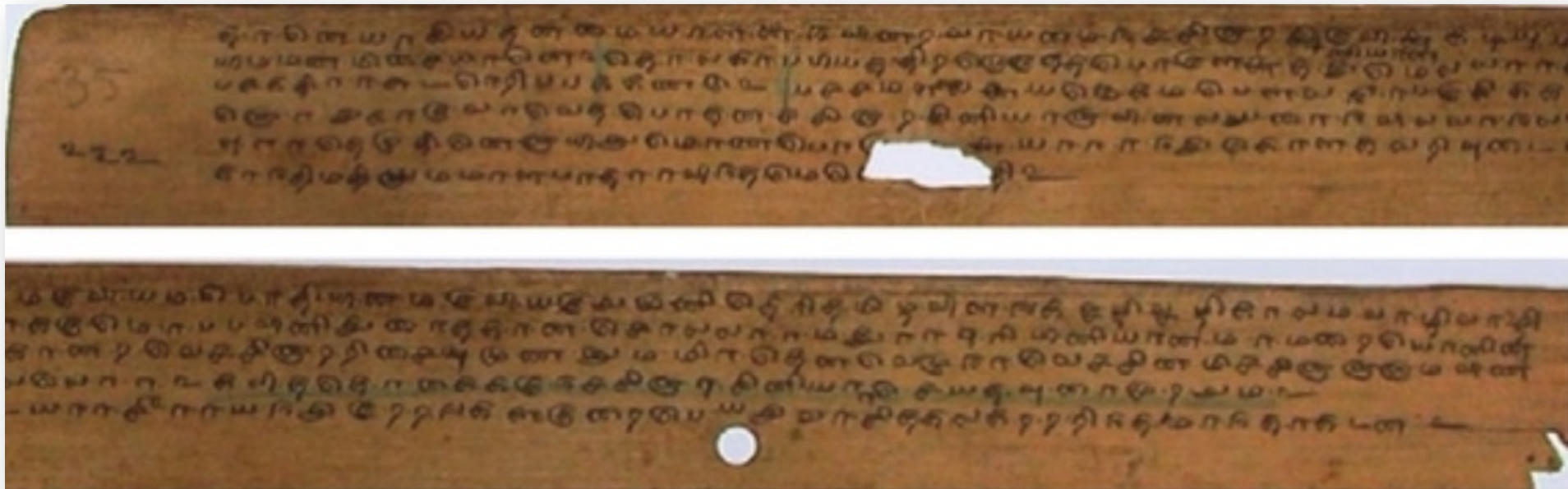
UNIVERSITÄT ZU LÜBECK

WWU MÜNSTER

AGENDA

1. Introduction to Semantic Systems [Tanya]

2. Supervised Learning [Marcel]

3. Unsupervised and Relational Learning [Magnus]

4. Summary [Tanya]

HELPING HUMANS:
TEXT UNDERSTANDING

- Picture by Eva Wilden, in: Tamil Satellite Stanzas: Genres and Distribution

TASKS: SYSTEM PERSPECTIVE

- Information retrieval can often be formulated as some form of *classification*

  - Part of text annotation or not?

  - Document relevant or not to a given search?

  - Which parts of a document are relevant?
    - To a given search string
    - For a summary

  - Exploration can include classification tasks but may also require different techniques

- How to realise a task depends, among other things, on which information is used from the documents
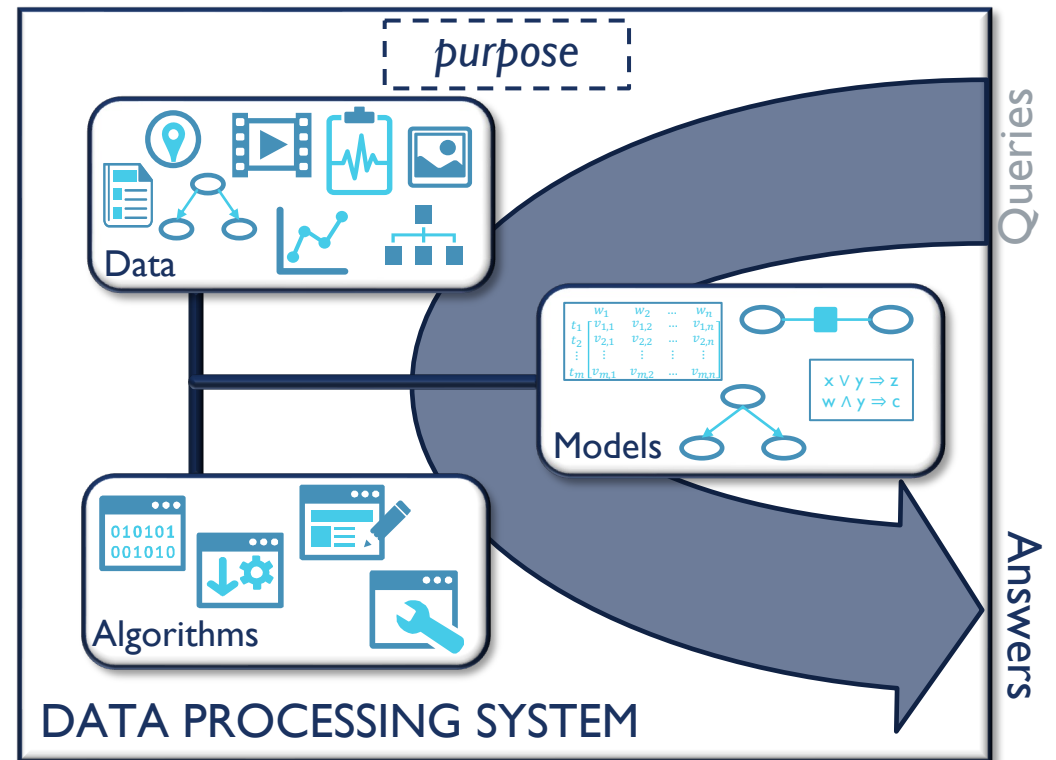
THE SETTING:
A CORPUS OF DOCUMENTS AND ANNOTATIONS

- Corpus = set of documents $\mathcal{D}$

- Each document d has a set of annotations $g(d)$

  - Annotation $\triangleq$ *subjective content description* (SCD)

  - Reflect the *context* of the purpose of the corpus

- Types of SCDs can be manifold

  - Figures, notes, references, …

Blei et al.

At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est.

<subject, predicate, object>

- Each SCD associated with words at specific locations throughout the corpus

  - Assumption: Words closer to location → influence higher

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Dolor sed viverra ipsum nunc aliquet bibendum enim. In massa tempor nec feugiat. Nunc aliquet bibend
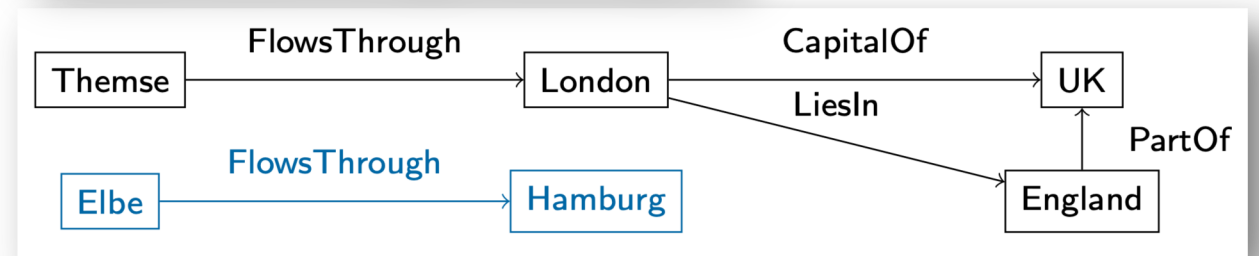
<SCD1>
<SCD2>
<SCD3>
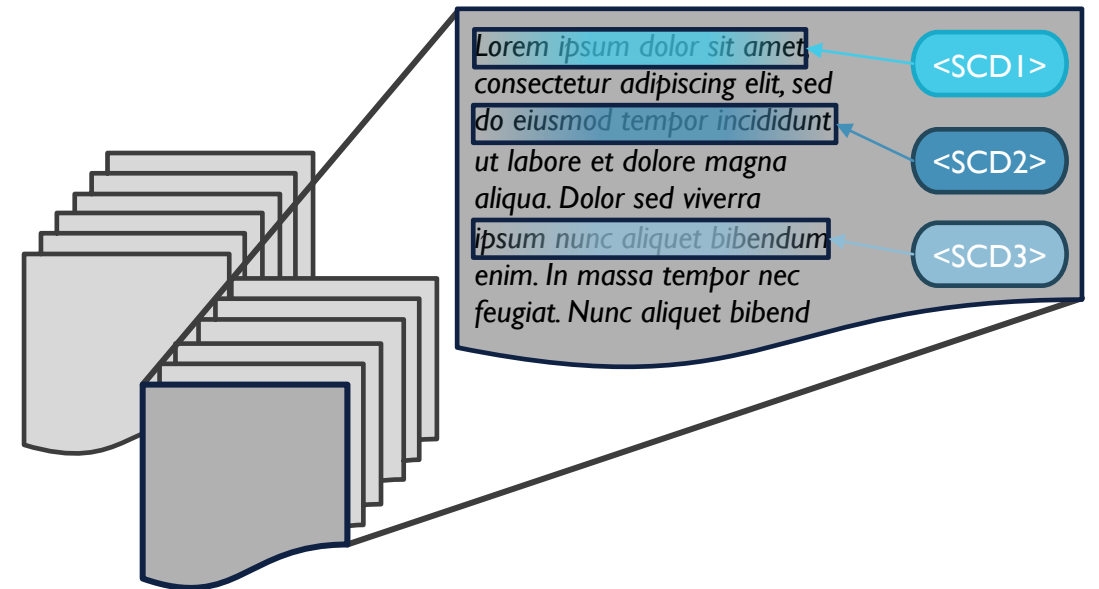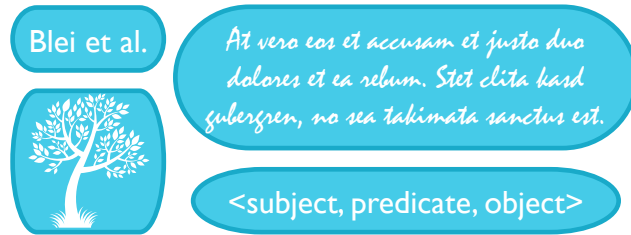
TOWARDS MORE EXPRESSIVE MODELS

- Hidden Markov model: special case of dynamic Bayesian network

  - *Probabilistic* → handle uncertainty / non-determinism

  - *Sequential* → handle behaviour over a sequence (time)

  - *Propositional* → handle features / attributes and classes

- But if you consider SPO triples, we have

  - Set of objects / individuals with attributes

  - Relations among them
    - See knowledge graphs

  and

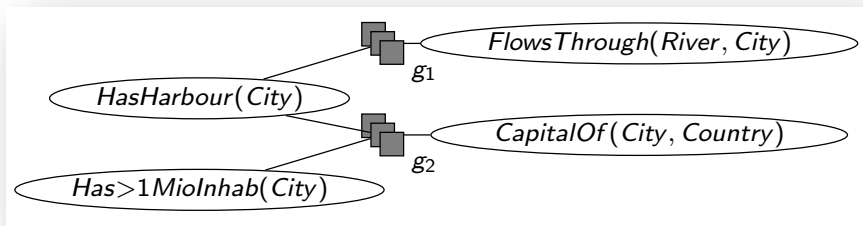 - *Uncertainty about existence, values, …*

<Themse, FlowsThrough, London>
<Elbe, FlowsThrough, Hamburg>
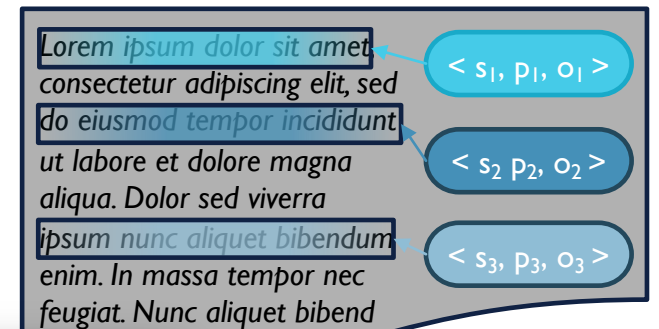
TOWARDS MORE EXPRESSIVE MODELS

■ Deal with worlds that are *uncertain* and have *things* in it!

- $FlowsThrough(River, City)$
- $HasHarbour(City)$ — g_1
- $CapitalOf(City, Country)$
- $Has>1MioInhab(City)$ — g_2

■ Use information about relations for

■ Compact encoding of an otherwise huge propositional model

■ Group (almost) indistinguishable instances

■ Efficient inference using representatives for groups

The world has things in it!

AI: intelligent systems in the real world

The world is uncertain!

First-order logic

Probabilistic graphical models

The world is uncertain!

Statistical Relational Artificial Intelligence
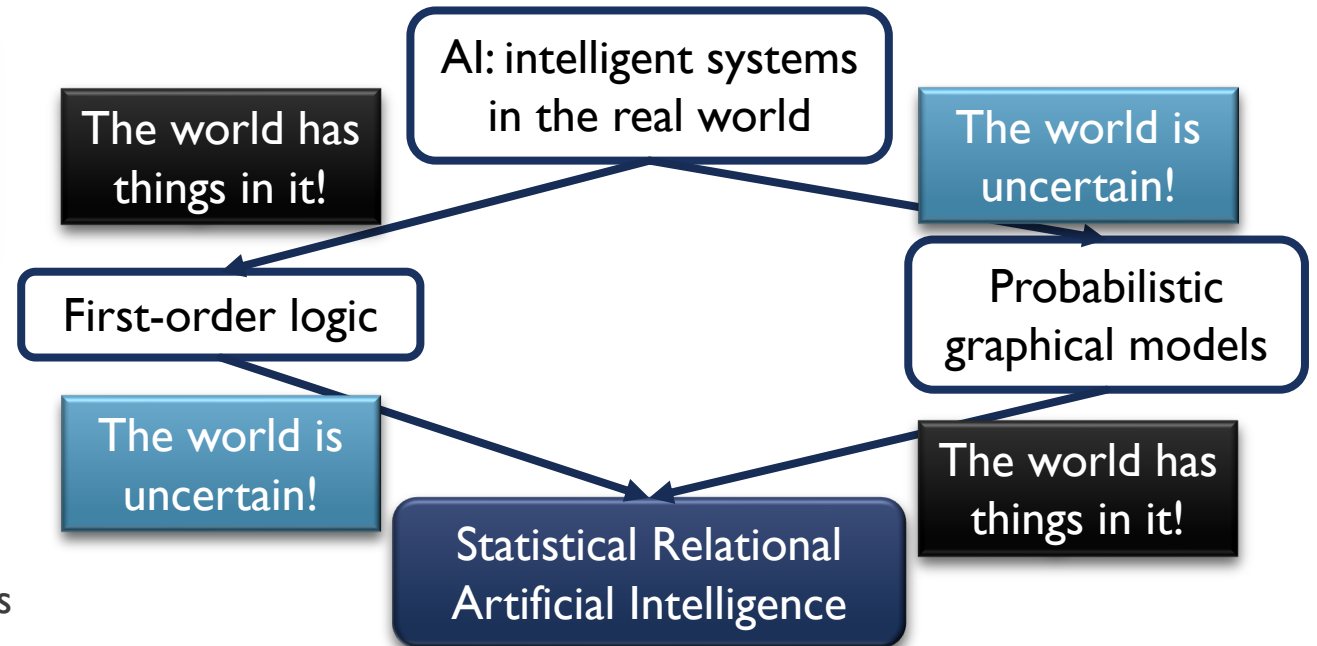
The world has things in it!
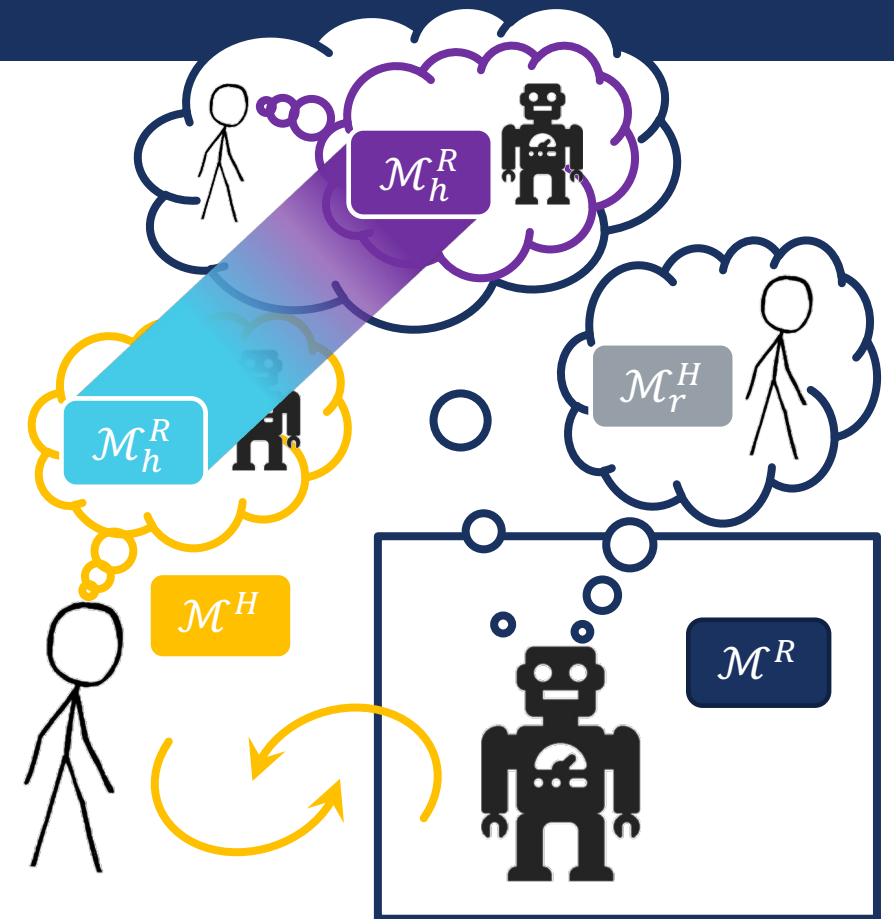
Figure based on Stuart Russell

KEEPING SYSTEMS UP TO DATE: CONTINUOUS TRAINING

- A trained system is stuck in the year of its inception if it cannot be adapted

- Take a note from, e.g., belief revision
  [e.g., Peppas & Williams, 1995]

  - Update a knowledge base to change old belief to new beliefs based on new information

  - Revise a knowledge base to keep it consistent when presented with new information

  - Notion of *forgetting*: remove facts, relations from knowledge bases
    [e.g., Beierle et al., 2019]

- First work on updating SCD systems (presented today at CHAI workshop by Magnus)

  - Incorporate feedback regarding wrong SCDs into a system
    [Bender et al., 2023]

HUMAN-CENTRED SYSTEMS

- Bring the human back into the loop

    - Provide interpretable services

    - Incorporate feedback or allow human to train system

- Take, e.g., a note from human-aware planning and decision making
  [e.g., Sreedharan et al. 2022]

    - Interpretable planning: explicability, legibility, predictability

    - Explanations through model reconciliation: updating a human's knowledge base with a system's knowledge

        - Revising a human's belief (not the system's belief)

$\mathcal{M}_h^R$

$\mathcal{M}_r^H$

$\mathcal{M}_h^R$

$\mathcal{M}^H$

$\mathcal{M}^R$

Figure based on material by Subbarao Kambhampati
http://rakaposhi.eas.asu.edu

WHAT ELSE IS THERE TO DO?

Thank you for having us!

- Incorporate additional requirements such as

  - Privacy

  - Ethics

  - Robustness

  - Resilience

- And so much more…

Link to tutorial website (QR code points to this website)
https://www.uni-muenster.de/Informatik.AGBraun/en/research/tutorials/ki-23.html

REFERENCES

- **Beierle et al. (2019)**
  Christoph Beierle, Gabriele Kern-Isberner, Kai Sauerwald, Tanja Bock, and Marco Ragni: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In KI – Künstliche Intelligenz, 2019.

- **Bender et al. (2023)**
  Magnus Bender, Kira Schwandt, Ralf Möller, and Marcel Gehrke: FrESH – Feedback-reliant Enhancement of Subjective Content Descriptions by Humans. In *CHAI 2023 CEUR-Proceedings of the 3rd Workshop on Humanities-Centred AI*, 2023.

- **Peppas & Williams (1995)**
  Pavlos Peppas and Mary-Anne Williams: Constructive Modelings for Theory Change. In *Formal Logic*, 1995.

- **Sreedharan et al. (2022)**
  Sarath Sreedharan, Anagha Kulkarni, and Subbarao Kambhampati: Explainable Human-AI Interaction: A Planning Perspective. Springer, 2022.