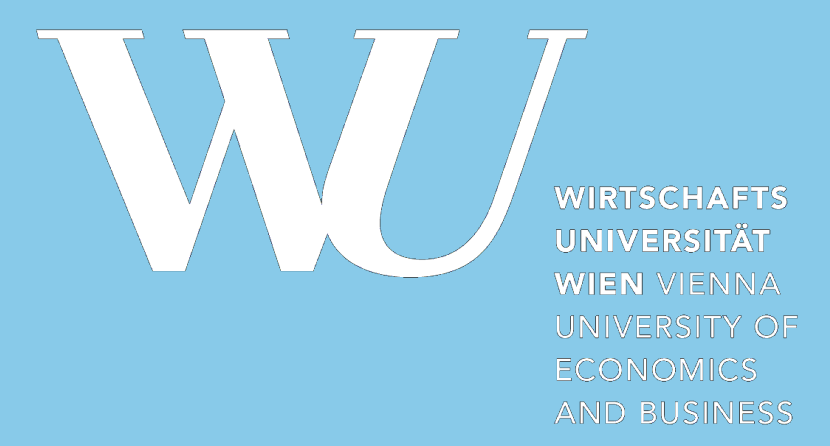
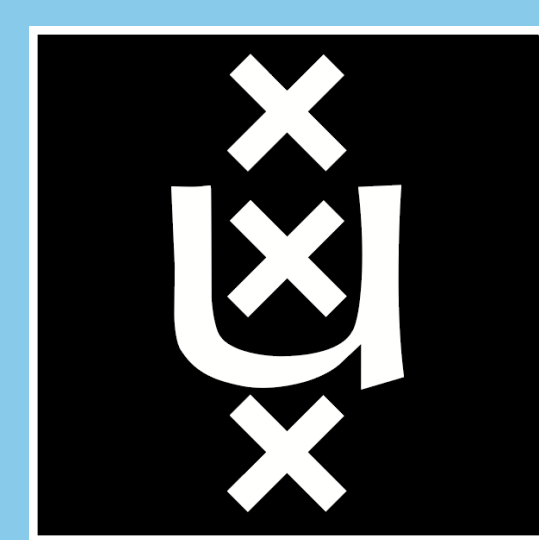


Conversational Exploratory Search via Interactive Storytelling

Svitlana Vakulenko*, Ilya Markov**, Maarten de Rijke**
 *WU Vienna - Institute for Information Business, Vienna, Austria
 **University of Amsterdam, Amsterdam, The Netherlands



MOTIVATION

1. Asking a good (well-formulated) questions is not easy
2. Question is **part of** the answer (reduces the search space)
3. Do not expect simple answers to complex questions (**beyond look-up queries**)

INTERACTIVE STORYTELLING

Computational narrative intelligence - ability to craft, tell, understand, and respond affectively to stories [1]

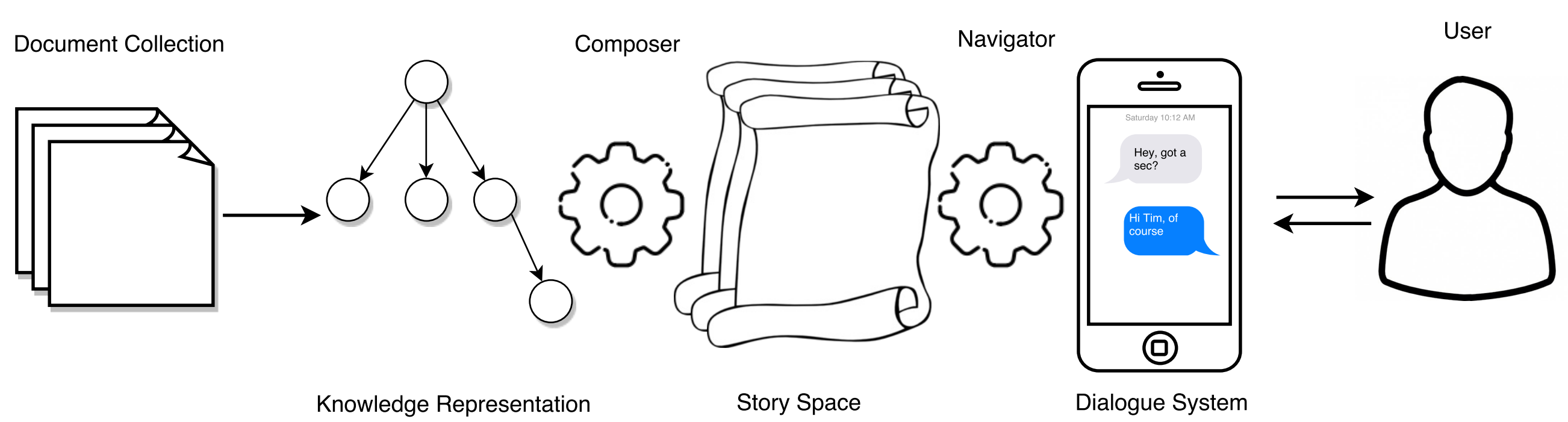


Figure 1. Communicating knowledge via an interactive storytelling process

RESEARCH QUESTION

1. How to make **expressive dialogs**?
i.e beyond simple QA
2. How to make **interactive stories**?
i.e. conversational story space navigation

STORY GENERATION

[3] Andrew Lang **fairy tales**: 437 stories

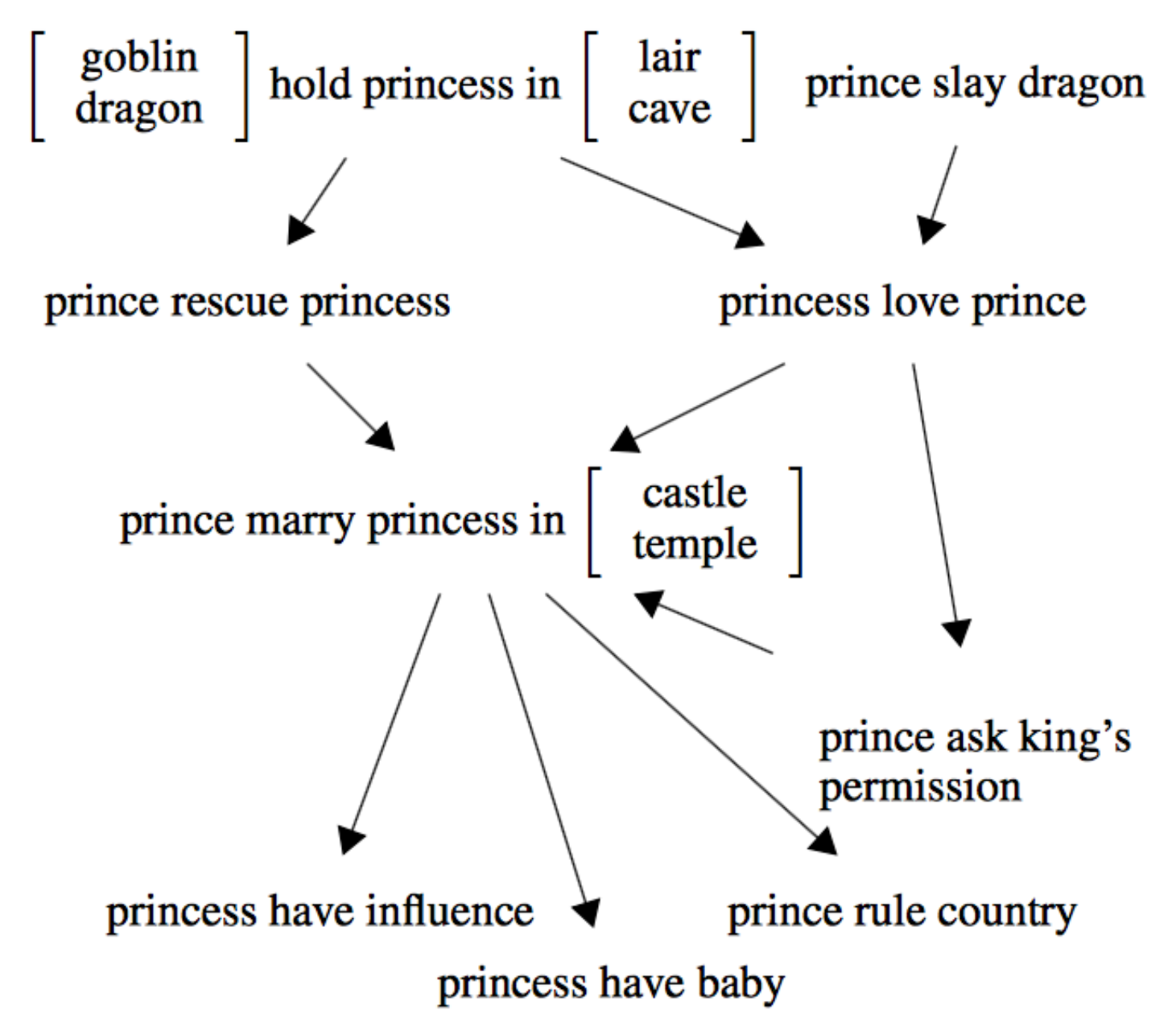


Figure 3: Plot graph

DIALOG STATE TRACKING [2]

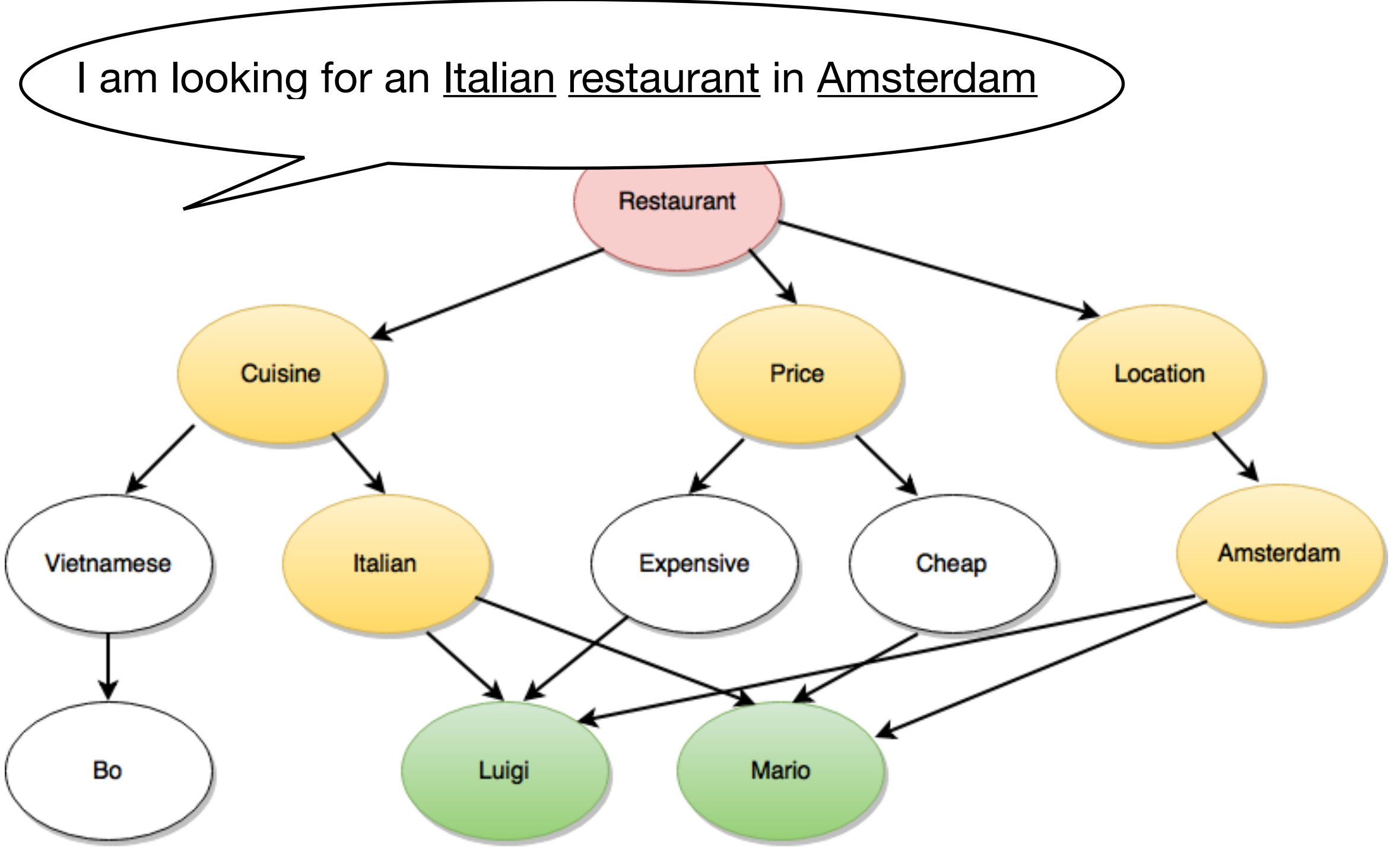


Figure 2: Domain ontology

[4] Wikipedia **movie plots**: 42,170 stories

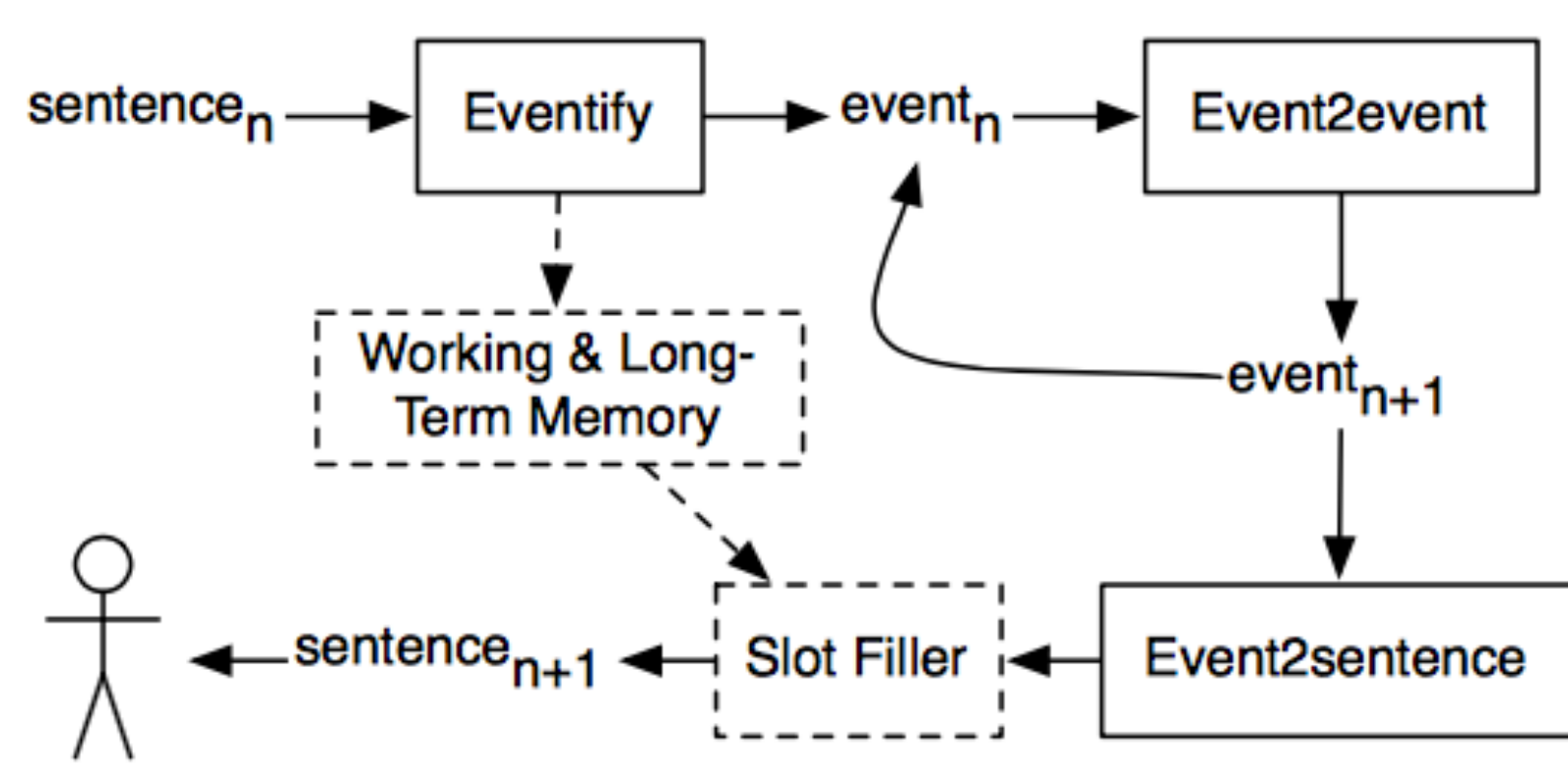


Figure 4: two Seq2Seq neural networks pipeline

REFERENCES

[1] Mark O. Riedl. Computational Narrative Intelligence: A Human-Centered Goal for Artificial Intelligence. CoRR abs/1602.06484 2016.
 [2] Jason D. Williams, Antoine Raux, and Matthew Henderson. 2016. The Dialog State Tracking Challenge Series: A Review. D&D 7, 3 (2016), 4–33.
 [3] Neil Duncan McIntyre and Mirella Lapata. Plot Induction and Evolutionary Search for Story Generation. ACL 2010.
 [4] Lara J. Martin, Prithviraj Ammanabrolu, William Hancock, Shruti Singh, Brent Harrison, and Mark O. Riedl. Event Representations for Automated Story Generation with Deep Neural Nets. CoRR abs/1706.01331 2017.

EVALUATION

- [3] Rate the stories on a scale of 1 to 5 for:
- **fluency** (was the sentence grammatical?),
 - **coherence** (does the story make sense?)
 - **interest** (how interesting is the story?)
- [4] Unigram **perplexity** ability of the model to predict the test data.

BLEU score considers all n-gram overlaps between the generated and expected output, where n varies from 1 to 4.

