Chapter NLP:I

Examples of NLP Systems

Question Answering: IBM Watson at Jeopardy

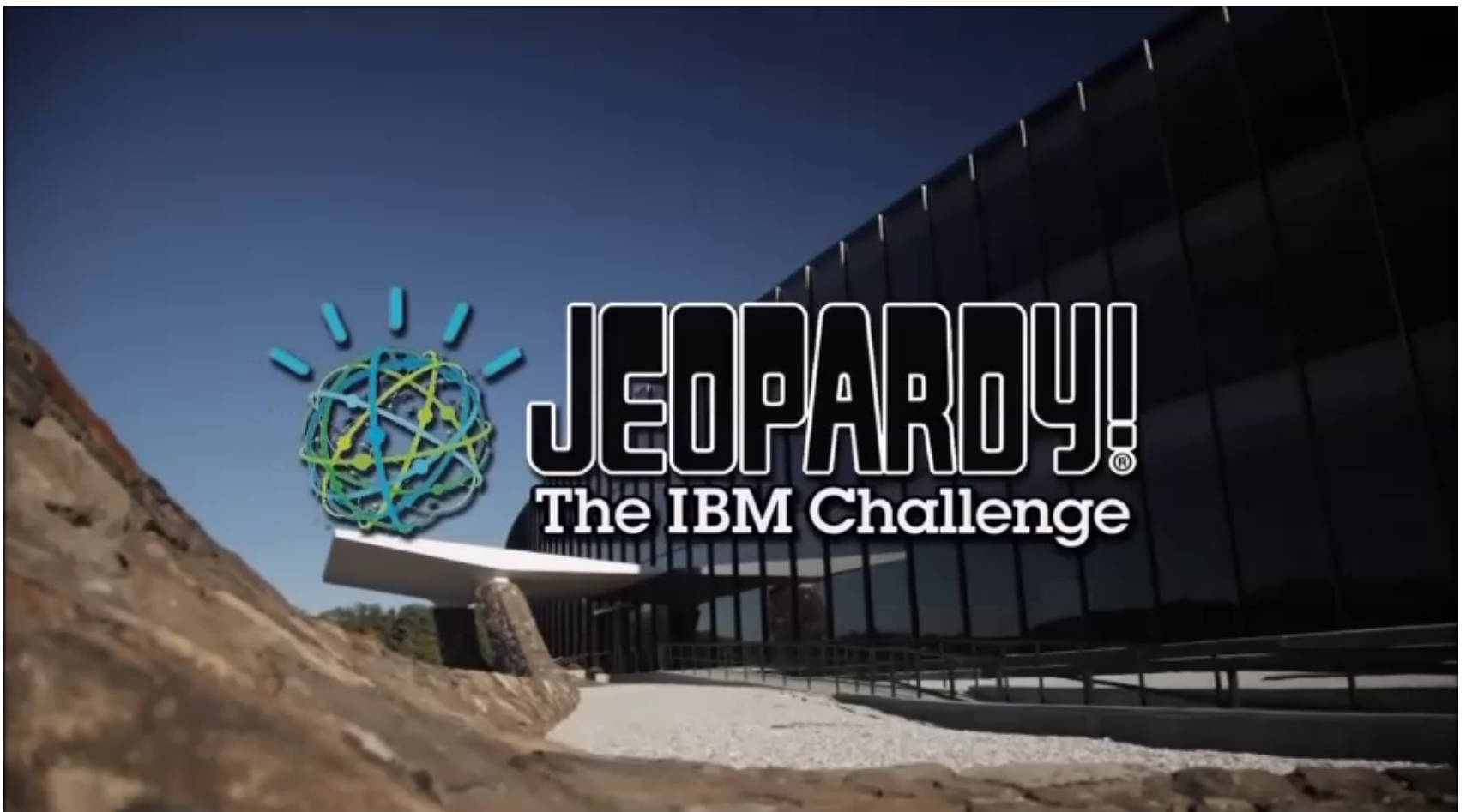
Jeopardy!

- American television quiz show running since the 1960s
- several general knowledge topics (e.g. history, literature, popular culture) at different dollar values
- participants presented with *clues in the form of answers*
- must formulate their *responses in the form of questions*

- between the 1960s and 2011 several returning champions; among others, Rutter and Jennings
- 2011: Rutter and Jennings vs. 200 million pages of content + AI (structured and unstructured, including full 2011 Wikipedia; ca. 4Tb of storage)

Examples of NLP Systems

Question Answering: IBM Watson at Jeopardy (continued)

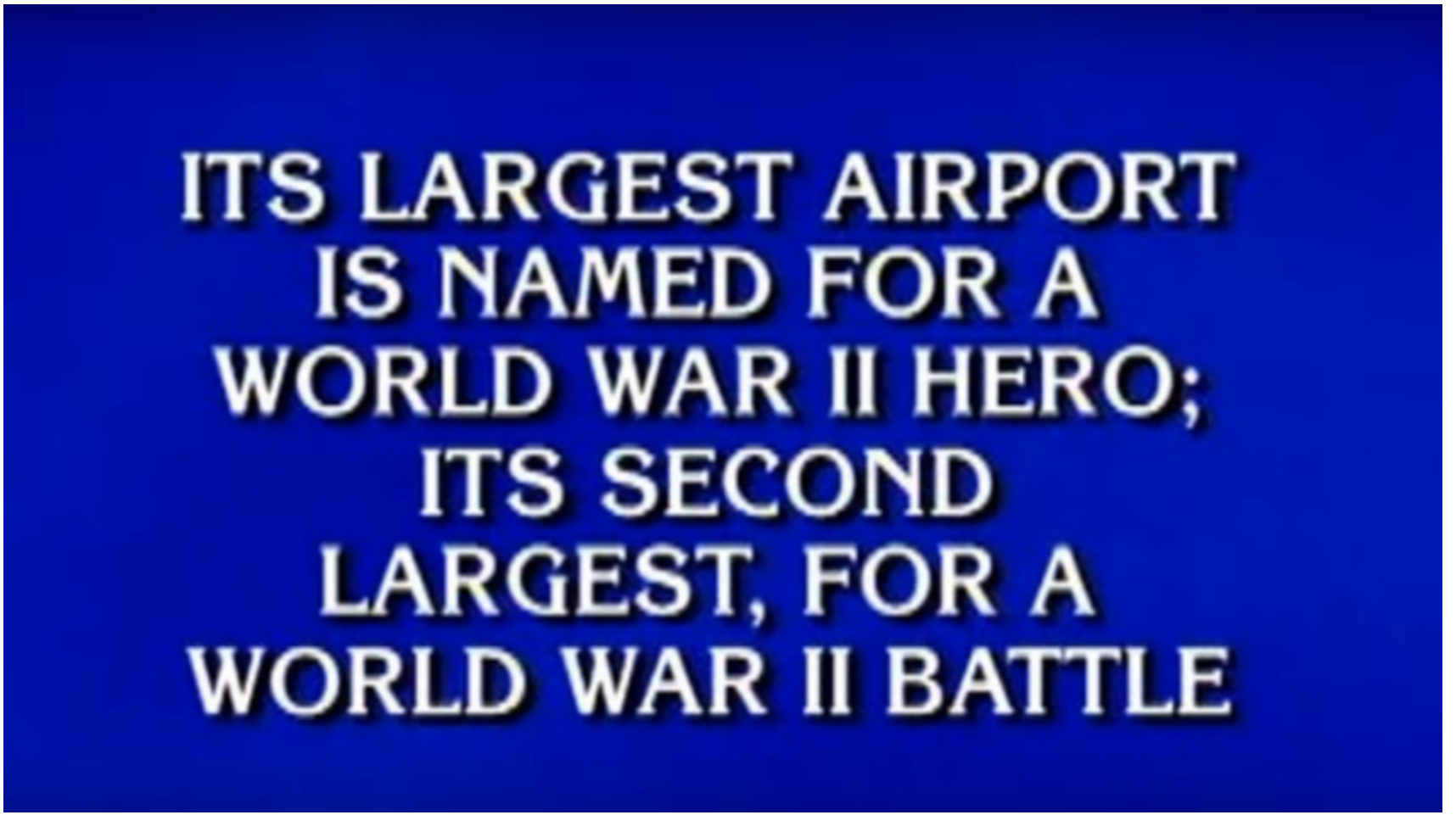


Examples of NLP Systems

Question Answering: IBM Watson at Jeopardy (continued)

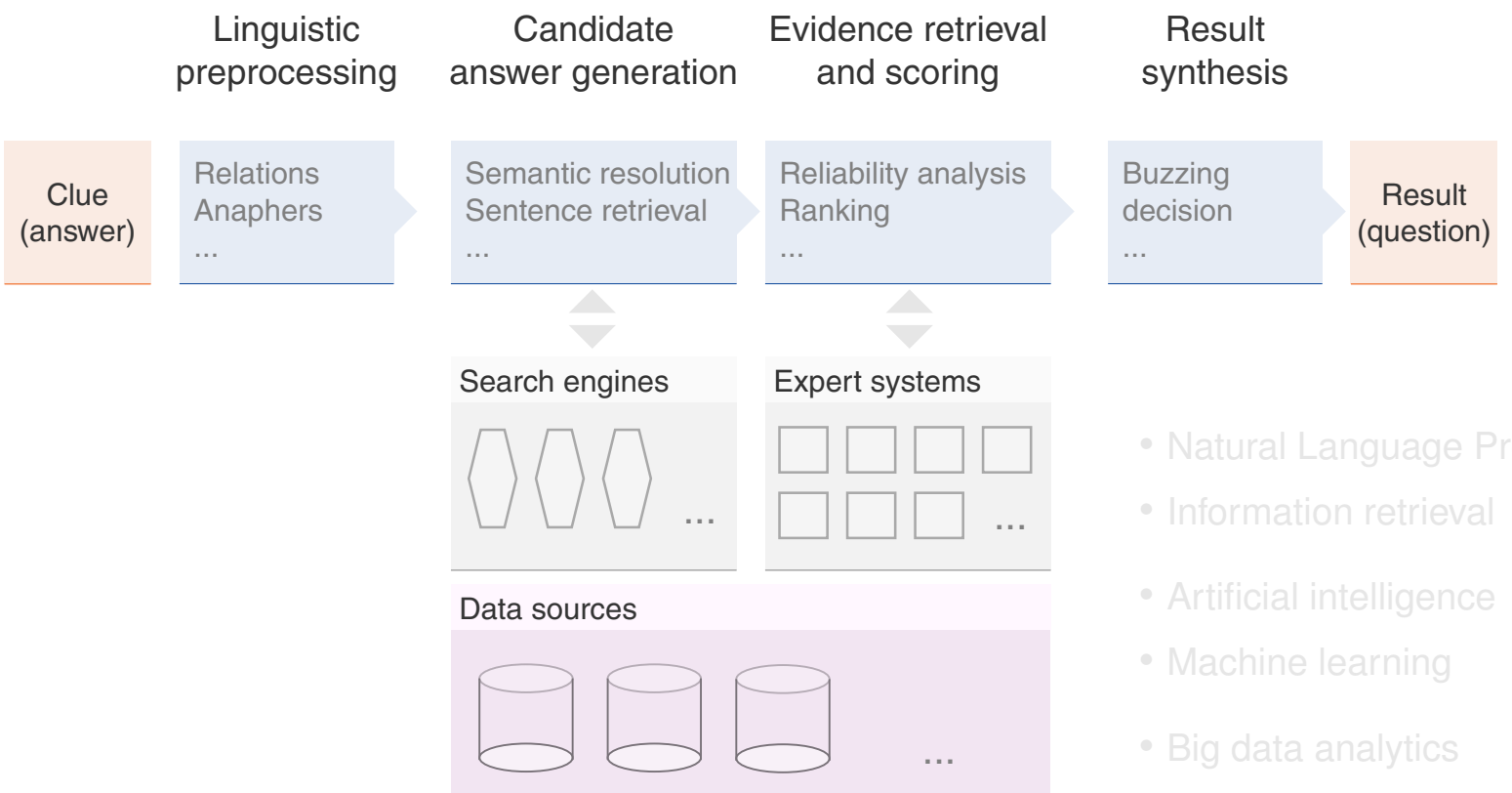
Examples of NLP Systems

Question Answering: IBM Watson at Jeopardy (continued)



Examples of NLP Systems

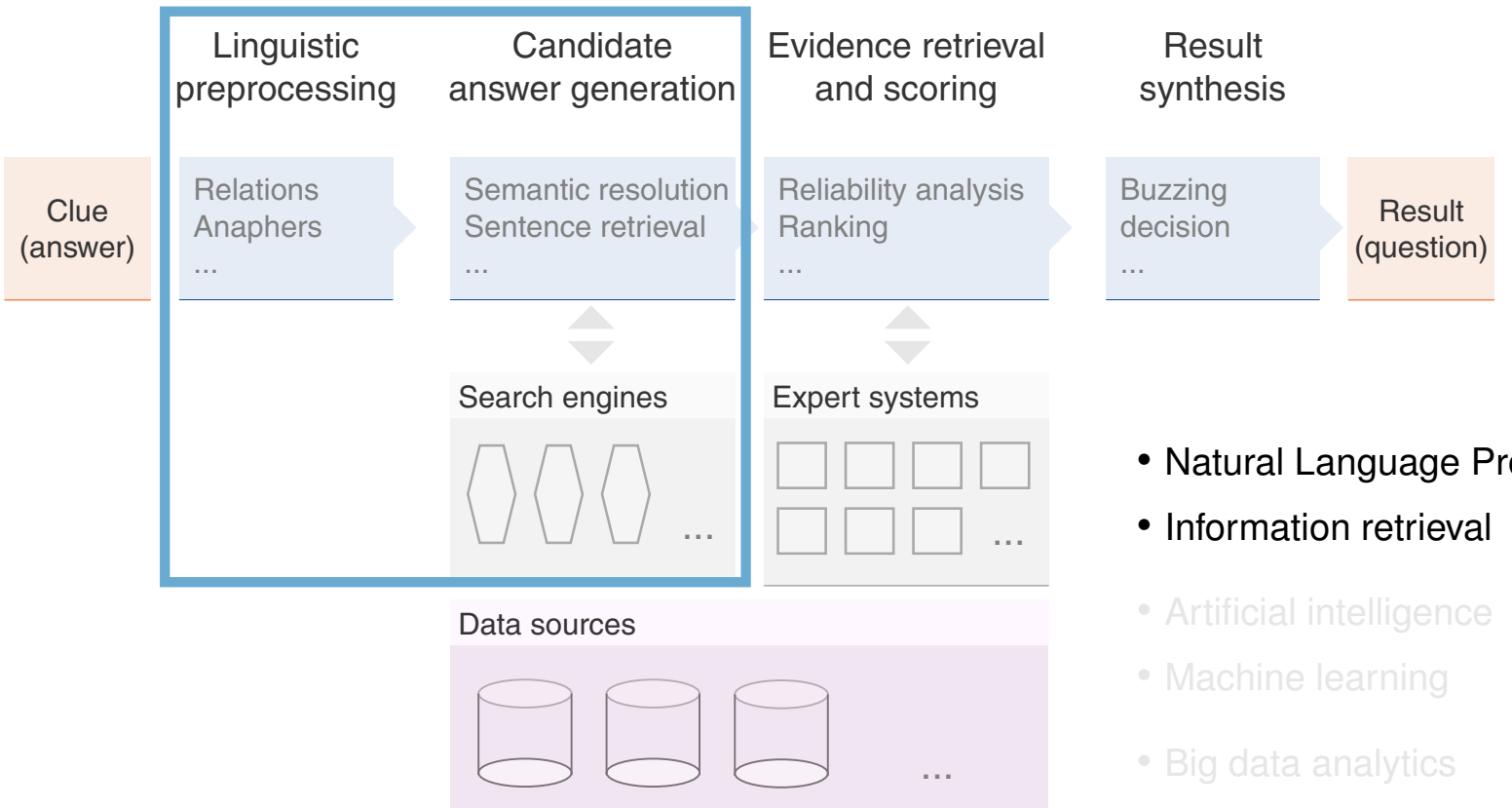
Question Answering: IBM Watson at Jeopardy (continued)



- Natural Language Processing
- Information retrieval
- Artificial intelligence
- Machine learning
- Big data analytics

Examples of NLP Systems

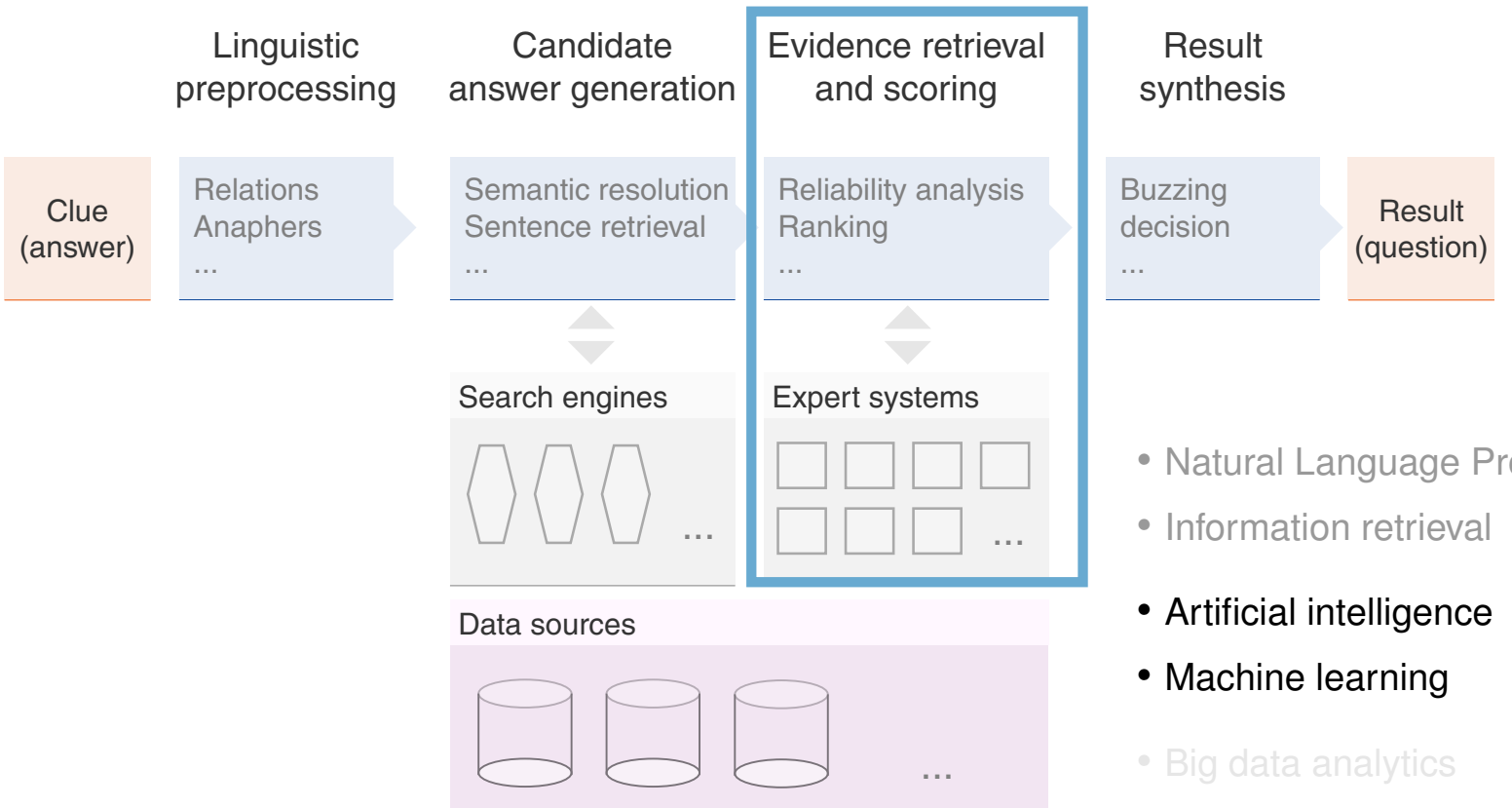
Question Answering: IBM Watson at Jeopardy (continued)



- Natural Language Processing
- Information retrieval
- Artificial intelligence
- Machine learning
- Big data analytics

Examples of NLP Systems

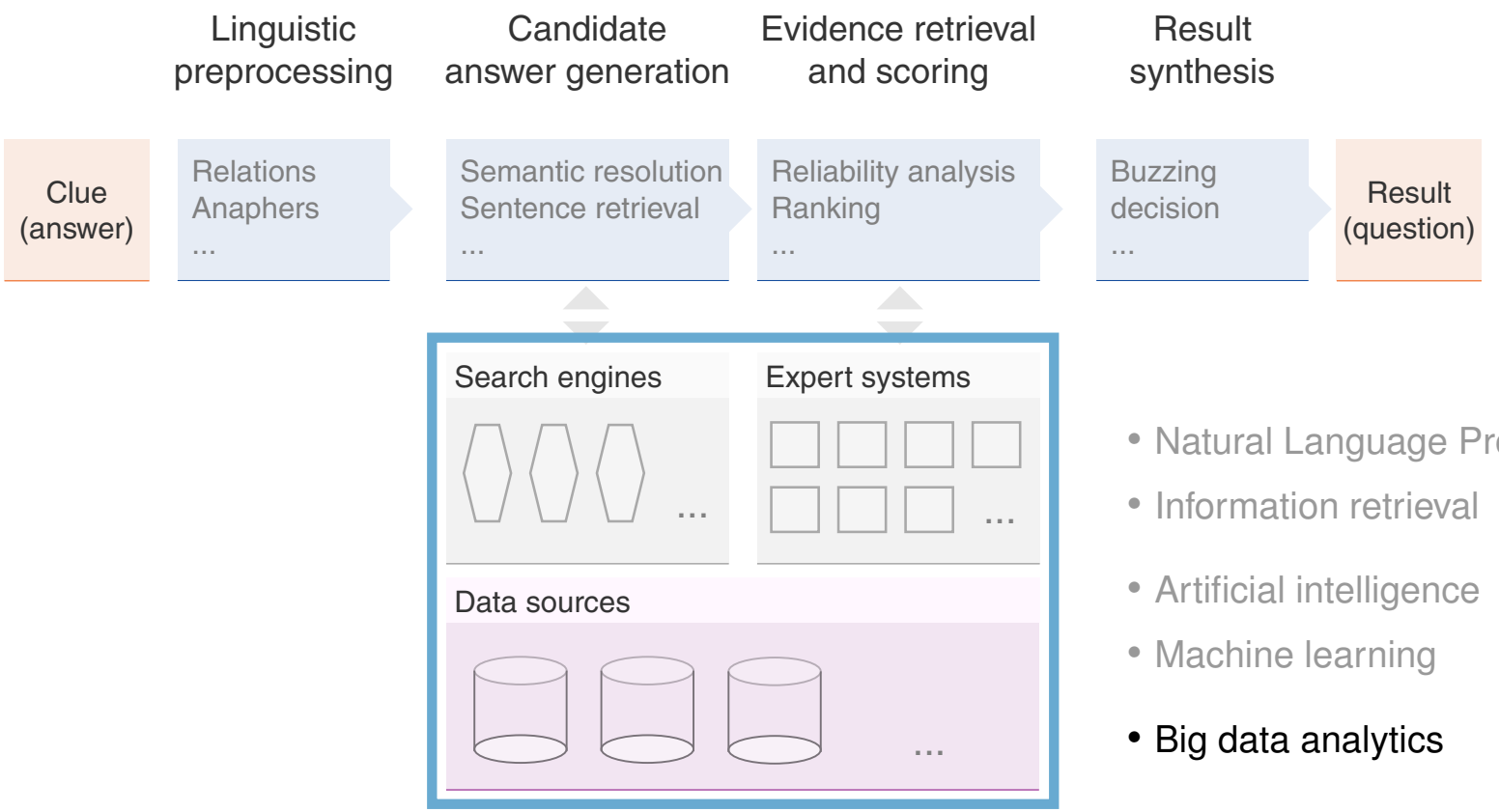
Question Answering: IBM Watson at Jeopardy (continued)



- Natural Language Processing
- Information retrieval
- Artificial intelligence
- Machine learning
- Big data analytics

Examples of NLP Systems

Question Answering: IBM Watson at Jeopardy (continued)



- Natural Language Processing
- Information retrieval
- Artificial intelligence
- Machine learning
- Big data analytics

NLP Problems

Information Extraction (IE)

NLP Problems

Review Analysis

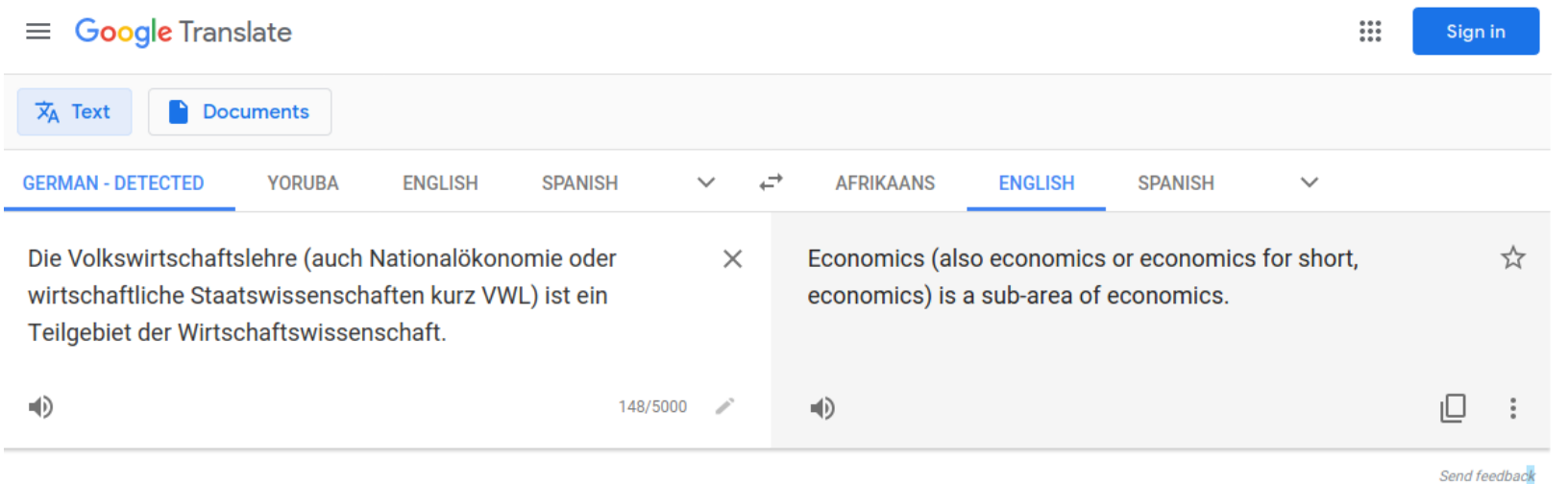
Attributes: zoom, affordability, size and weight, flash, ease of use

Size and weight:

- ✓ Nice and compact to carry!
- ✓ Since the camera is small and light, I won't need to carry around those heavy, bulky professional cameras either!
- ✗ The camera feels flimsy, is plastic and very light in weight you have to be very delicate in the handling of this camera

NLP Problems

Machine Translation (MT)



First sentence of the Wikipedia article on “Volkswirtschaftslehre”.

See also twitter.com/hashtag/googletranslatefails

NLP Problems

State of Affairs: Mostly Solved

- Spam detection.
  Let's go to Agra vs. Buy V1Agra

- Part-of-speech (POS) tagging.
  Colorless/Adjective green/Adjective ideas/Noun sleep/Verb furiously/Adverb.

- Named entity recognition (NER).
  Einstein:Person met with UN:Organization officials in Princeton:Location.

NLP Problems

State of Affairs: Making Good Progress

- Sentiment detection.
  Best pizza in town. vs. The waiter ignored us for 20 minutes.

- Coreference resolution.
  ?My trophy did not fit into ?the suitcase because it is too big.

- Word sense disambiguation (WSD)
  I need new batteries for my mouse.

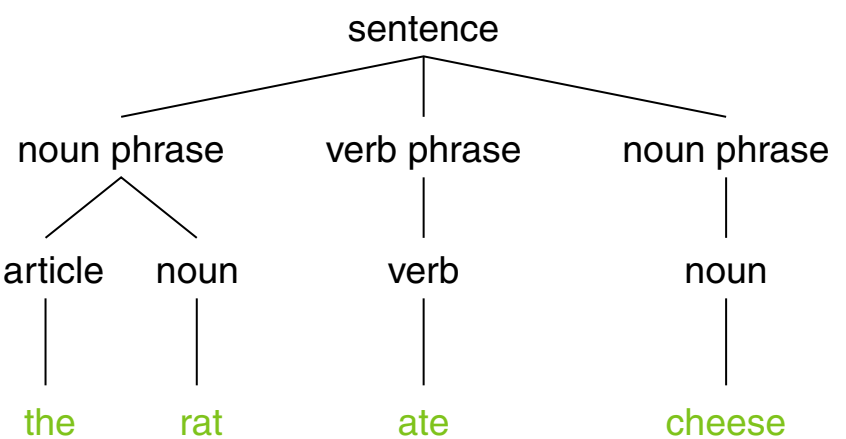
NLP Problems

State of Affairs: Making Good Progress (continued)

- Machine translation.
  Is getting better and better. → Wird immer besser.

- Information extraction.
  Come to our first lecture, April 3. → Calendar update: Lecture (April 3)

- Parsing.
  The rat ate cheese. →



NLP Problems

State of Affairs: Still Challenging

- Question answering (QA).
  Is ibuprofen effective in reducing fever for patients with acute febrile illness?
- Paraphrasing.
  XYZ acquired ABC yesterday vs. ABC has been taken over by XYZ
- Summarization.
  Dow Jones is up + house prices rose $\rightarrow$ Economy is good
- Dialogue.
  User: Best pizza around?
  Echo/Siri/Now: Antonio's. Want a table tonight?

Remarks:

- On referring to the field (roughly):
  - Natural Language Processing/Language Engineering: devising methods for processing specific language phenomena (e.g. resolving pronouns); operationalizing formal models of language (e.g. computational formal grammars)
  - Language Technology/Text Technology/Speech Technology: applications of NLP (various sub-areas: MT, Dialogue Systems, etc.)
  - Computational Linguistics: Linguistics/Language science research using computational means

  These terms are (unfortunately) often used interchangeably.

- For an overview of history of NLP see, for example, Karen Sparck Jones (1994) Natural Language Processing: A Historical Review

- Food for thought: 2019 IBM Project Debater held its first public live debate with Harish Natarajan who holds the world record for most debate competitions won; the event can be viewed here. Watch (parts of) the debate and then go back to the schema of Watson's architecture.
  - What kind of functionalities/functional components do you think are required for such a system?
  - Can you decompose the debating task into components, some of which require NLP?