Artificial Intelligence and the Future of Information Organization and Retrieval

Yaşar Tonta

Hacettepe University

Department of Information Management

X: @yasartonta

Email: yasartonta [at] gmail.com

Web: https://yunus.hacettepe.edu.tr/~tonta/

"Libraries as Transformers," IATUL Fall Seminar, 4-7 September 2023, İzmir National Library, İzmir

Enter search terms:
Search
in this collection
Advanced Search
Notify me via email or RSS

Links for Authors
Policies and Help Documentation

Browse
Collections
Disciplines

2003 IATUL PROCEEDINGS

The International Association of Scientific and Technological University Libraries (IATUL) is a voluntary international organization which provides a forum for the exchange of ideas relevant to librarianship in technological universities throughout the world. It also provides library directors and senior managers an opportunity to develop a collaborative approach to solving common problems.

The 24th Annual IATUL Conference was held on June 2nd – June 5th, 2003 at the Middle Eastern Technical University. The 24th Annual IATUL Conference's theme was "Libraries and Education in the Networked Information".

The final conference program for the conference can be found at this link: 2003 IATUL Conference Program.

Personalization of Digital Information Services

Yasar Tonta, *Hacettepe University Department of Information Management*

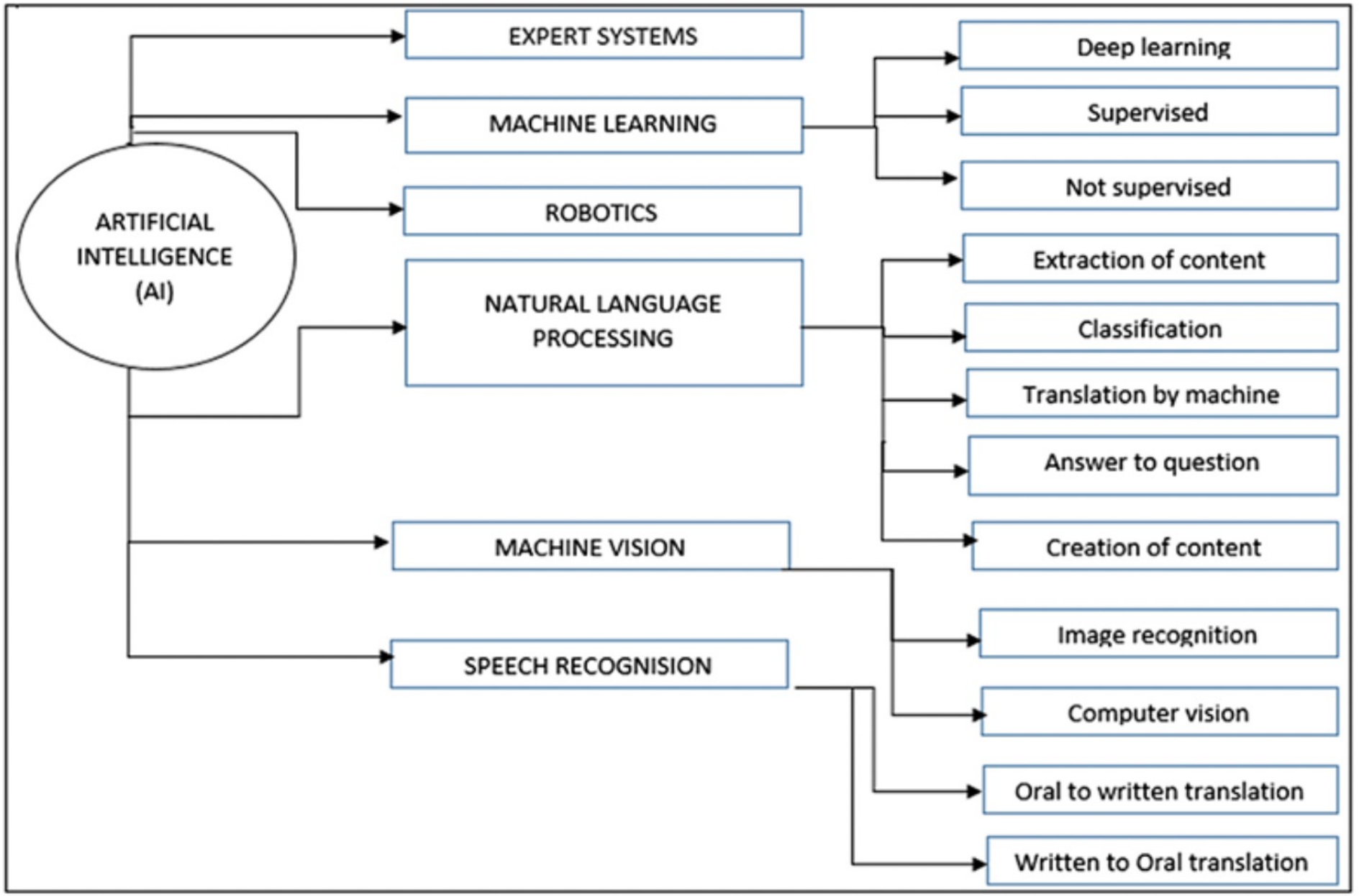
Plan

- Artificial Intelligence (AI)
  - Recursive Neural Networks (RNNs)
  - Large Language Models (LLMS)
- AI Techniques and Applications
- Search Engines vs. Conversational Chatbots
- Information Search and Retrieval
- Conversion of Search Engines to Question Answering (Q/A) Systems
- Can AI-based Chatbots Become Q/A Systems in the Future?
- Challenges Related to AI in Information Retrieval

Artificial Intelligence (AI)

- Recursive Neural Networks (RNN)
- Large Language Models (LLM)

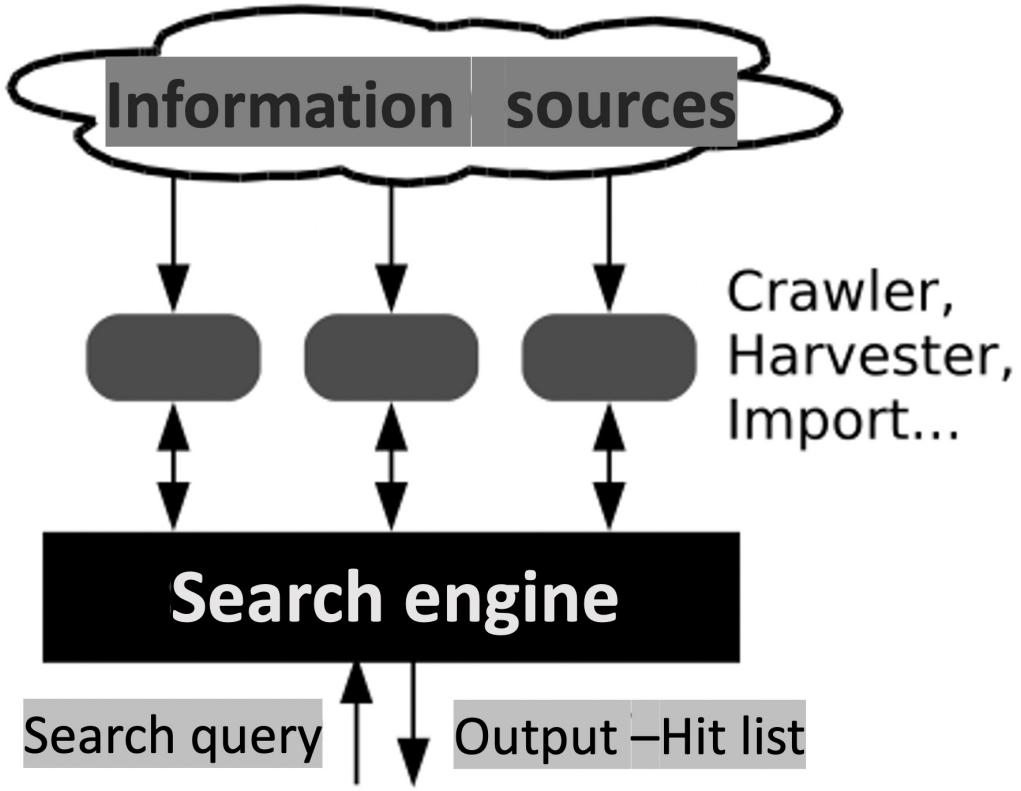
AI TECHNIQUES AND APPLICATIONS



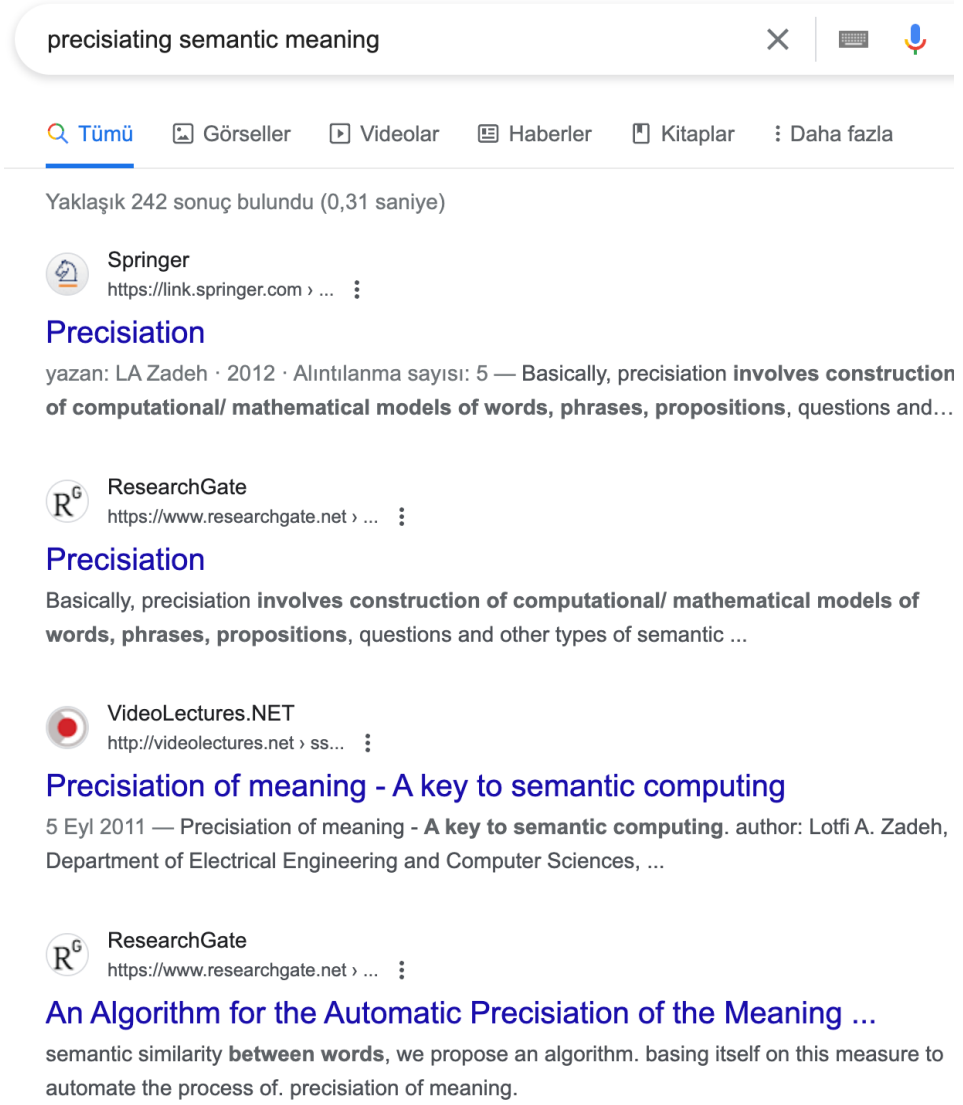
Source: Okunlaya et al. (2022)

Search Engines vs. Conversational Chatbots -I

- Search engines



Source: https://de.wikipedia.org/wiki/Suchmaschine



Search Engines vs. Conversational Chatbots -II

- Conversational chatbots (e.g., ChatGPT)

How does ChatGPT precisiate the semantic meaning?

ChatGPT uses a combination of statistical and semantic analysis, as well as machine learning algorithms, to precisiate the semantic meaning of the text it encounters.

At a high level, ChatGPT works by analyzing large amounts of text data and identifying patterns and relationships between different words and phrases. It does this by representing each word as a vector in a high-dimensional space, where words that are semantically similar are located close together. This process is known as word embedding, and it allows ChatGPT to capture the semantic meaning of words in a way that is independent of their surface form.

Source: ChatGPT: https://chat.openai.com/c/dd486cee-5007-42db-a977-c14e6fbcb5c3

Does That Mean It Is the End of the Hit Lists?

- Probably not. At least, not yet.
- Chatbots sometimes become too “creative”
- The source of information cannot easily be verified
- Information provided may not be up-to-date
- Answers provided by chatbots might be biased for various reasons (e.g., size of trained data sets, lack of scholarly content)
- . . .

Source: Bager (2023)

Besides, there are more fundamental issues . . .

Information Search and Retrieval

An ideal information retrieval (IR) system . . .

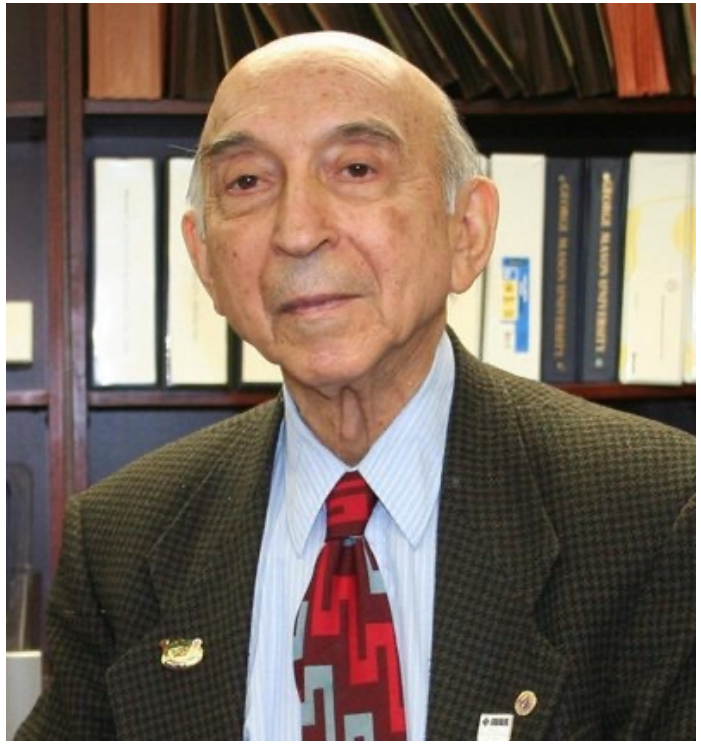
- Retrieves **all** and **only** relevant documents
  - Needle(s) in the haystack problem
- "Relevance": A measure of similarity between a search query and a document
- The main paradox of IR: "The need to describe that which you do not know in order to find it" (Roland Hjerrpe)
- Cf. "prompt engineering" in chatbots

Language

- Language both enables and limits our ability to communicate with IR systems
- Is “tomato” a fruit or vegetable?
  - “ **The age-old question actually has an answer—it's both!** Tomatoes are fruits that are considered vegetables by nutritionists.” (*Britannica*, https://www.britannica.com/story/is-a-tomato-a-fruit-or-a-vegetable
- Under which subject would you classify “Alternative Medicine” (or “Complementary Medicine”)?
  - Medicine?
  - Religion?
  - Philosophy?

Lotfi Zadeh (1921-2017)

Founder of fuzzy logic; soft computing; computing with words (CWW); and human level artificial intelligence



Source: https://www.azernews.az/culture/49308.html

Search Engines According to Lotfi Zadeh

- insufficient (they lack world's knowledge)
- work mostly with two-valued logic (relevant/not relevant)
- can't make inferences

Source: Zadeh (2005, 2006)

Conversion of Search Engines to Question Answering (Q/A) Systems

- World's knowledge
- Relevance (statistical/semantic)
  - q: How old is Vera?
  - p: Vera is the same age as Irene
  - r: Irene is 65
- Making inferences from perception-based knowledge
  - 2-valued logic and probability is not valid
- The main problem of search engines is understanding natural language
- The precisiation of meaning

Source: Zadeh (2005, 2006, 2011, 2012)

Can AI-based Chatbots Become Q/A Systems in the Future?

- As of now, AI
  - has yet to have the world's knowledge
  - does not quite "understand" the natural language
  - cannot precisiate the meaning
- So, AI does not seem to be aware of what it is doing.

- On the other hand . . .

  "No computer has ever been designed that is ever aware of what it's doing; but most of the time, we aren't either."

  --Marvin Minsky (the father of AI)

Challenges Related to AI in Information Retrieval

- Semantic computing
- Semantic search engines and/or AI-based Q/A chatbots
- Quantum IR (C.J.K. van Rijsbergen)
- Brain-computer interfaces
- Ethical and responsible use of AI

Sources

- Bager, J. (2023). Das Ende der Trefferlisten: Suchmaschinen beantworten Fragen mit Hilfe von KI. *Heise Magazine,* 20. https://www.heise.de/select/ct/2023/4/23006105559334242 40
- Dempsey, L. (2023, May 23). Generative AI and large language models: background and contexts. *LorcanDempsey.net,* https://www.lorcandempsey.net/intro-gen-ai/
- Okunlaya, R.O., Syed Abdullah, N. and Alias, R.A. (2022), Artificial intelligence (AI) library services innovative conceptual framework for the digital transformation of university education, *Library Hi Tech*, 40(6), 1869-1892. https://doi.org/10.1108/LHT-07-2021-0242
- Zadeh, L.A. (2005). From search engines to question-answering systems – The problems of world knowledge, relevance and deduction (keynote lecture). 6th WSEAS International Conference on Fuzzy Systems (FS '05), June 16-18, 2005, Lisbon, Portugal. http://www.worldses.org/plenary/2005/lisbon/zadeh-2005.pdf
- Zadeh, L.A. (2006). From search engines to question answering systems – The problems of world knowledge, relevance, deduction and precisiation. Sanchez, E. (ed). In *Fuzzy Logic and the Semantic Web* (p. 163-210). Amsterdam: Elsevier.
- Zadeh, L.A. (2011). Precisiation of Meaning: A Key to Semantic Computing. (SSSC, August 9, 2011).
- Zadeh, L.A. (2012). Computing with Words: Principle concepts and ideas. *Studies in Fuzziness and Soft Computing*, 277: 1-153.

Artificial Intelligence and the Future of Information Organization and Retrieval

Yaşar Tonta

Hacettepe University

Department of Information Management

X: @yasartonta

Email: yasartonta [at] gmail.com

Web: https://yunus.hacettepe.edu.tr/~tonta/

"Libraries as Transformers," IATUL Fall Seminar, 4-7 September 2023, İzmir National Library, İzmir