

Knowledge Discovery & Database Research Laboratory

데이터베이스 및 데이터마이닝 연구실

심규석 교수님



Degrees

Ph.D. 1993 Computer Science, University of Maryland, College Park
M.S. 1988 Computer Science, University of Maryland, College Park
B.S. 1986 서울대학교 전기공학과

Work Experiences

2002 - 현재 서울대학교 전기컴퓨터 공학부 교수
2001 - 2003 Microsoft Research, Visiting Scientist
1999 - 2001 카이스트 전산학과 (조)교수
1996 - 1999 Lucent Bell Laboratories, Member of Technical Staff
1994 - 1996 IBM Almaden Research Center, Research Staff
1993 - 1994 Federal Reserve Board, Research Staff
1992 - 1993 Hewlett-Packard Laboratories, Summer Intern

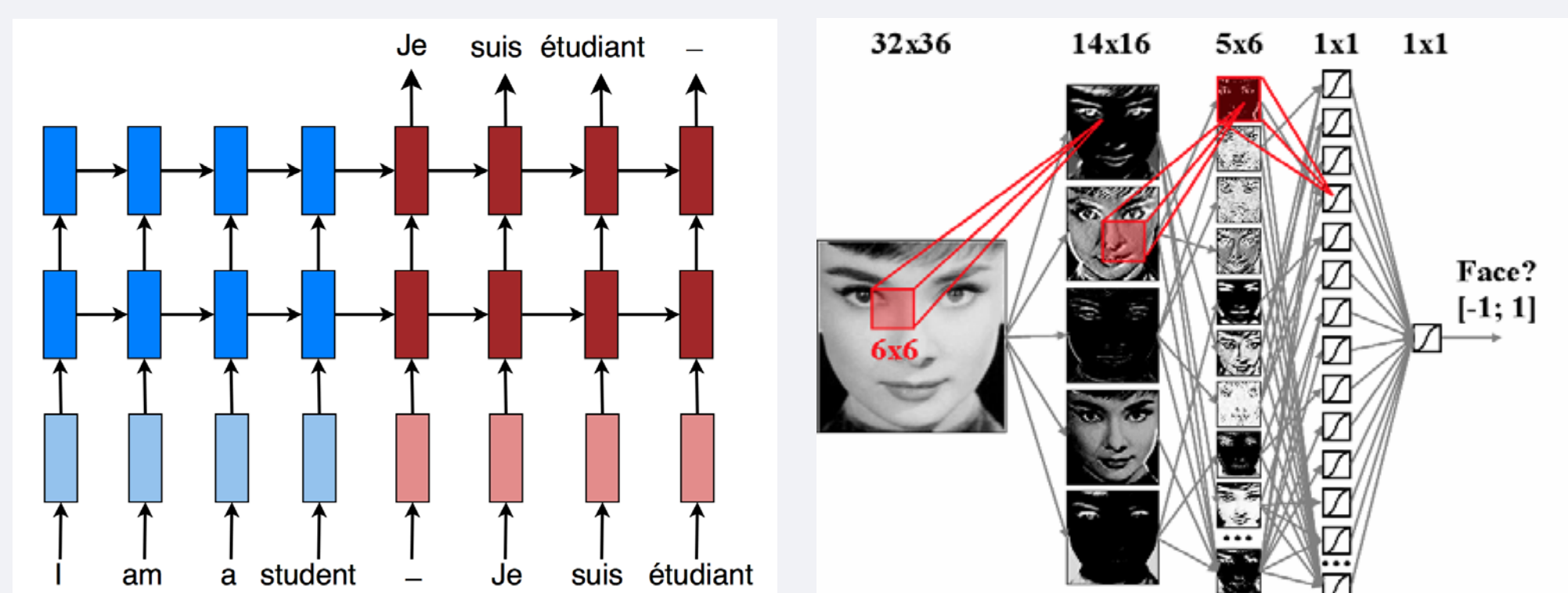
Research Interests

- Big Data Mining and Knowledge Discovery
- Deep Learning
- Query Processing and Optimization
- MapReduce Algorithms for Big Data (Parallel Algorithms)
- Recommendation Systems
- Data Privacy Preservation
- Histograms and Wavelet Synopsis
- Question Answering Systems
- Pattern Mining

Prerequisites

- Introduction to Data Structures
- Introduction to Algorithms

Deep Learning



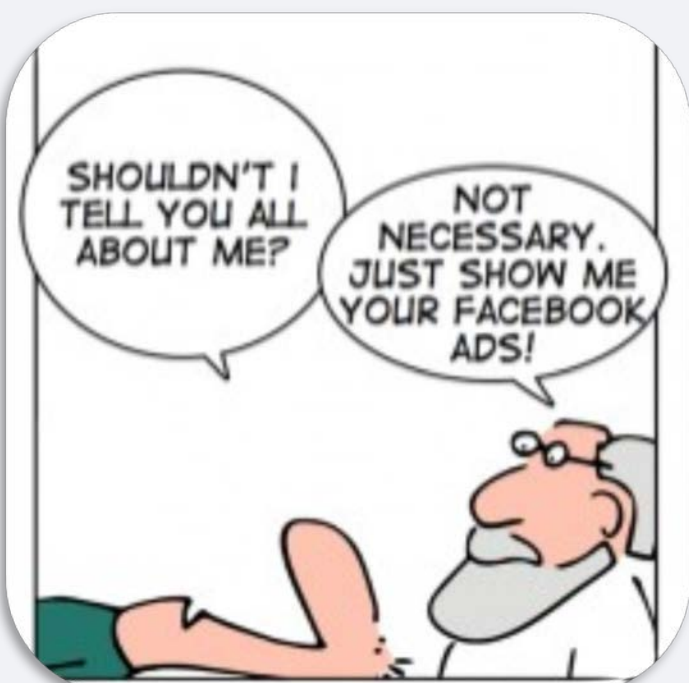
Deep learning architectures such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have been applied to fields including computer vision, speech recognition, natural language processing, machine translation and bioinformatics.

Recommendation Systems



Top-k followee recommendation: For each user, we choose top-k users whom the user would like to follow the most.
Top-k tweet recommendation: For each user, we find top-k tweets of others which the user would like to read the most.

Privacy Preservation



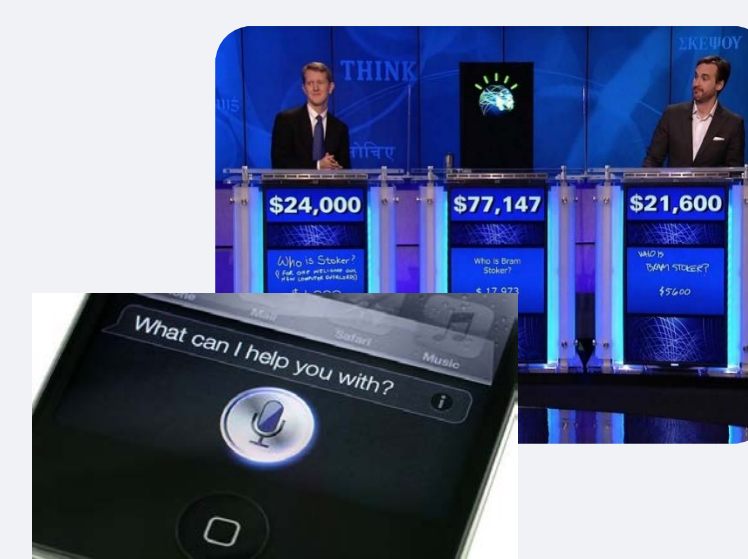
by Geek & Poke

For this reason, **privacy-preserving data mining** and **publishing** techniques are essential to protect privacy and exploit large amounts of data safely.

The Internet enables us to do lots of things such as sharing photos, shopping, banking, and so on. However, sensitive information accumulated on the web incurs many social problems and economic losses, attracting public attention to **privacy preservation**.

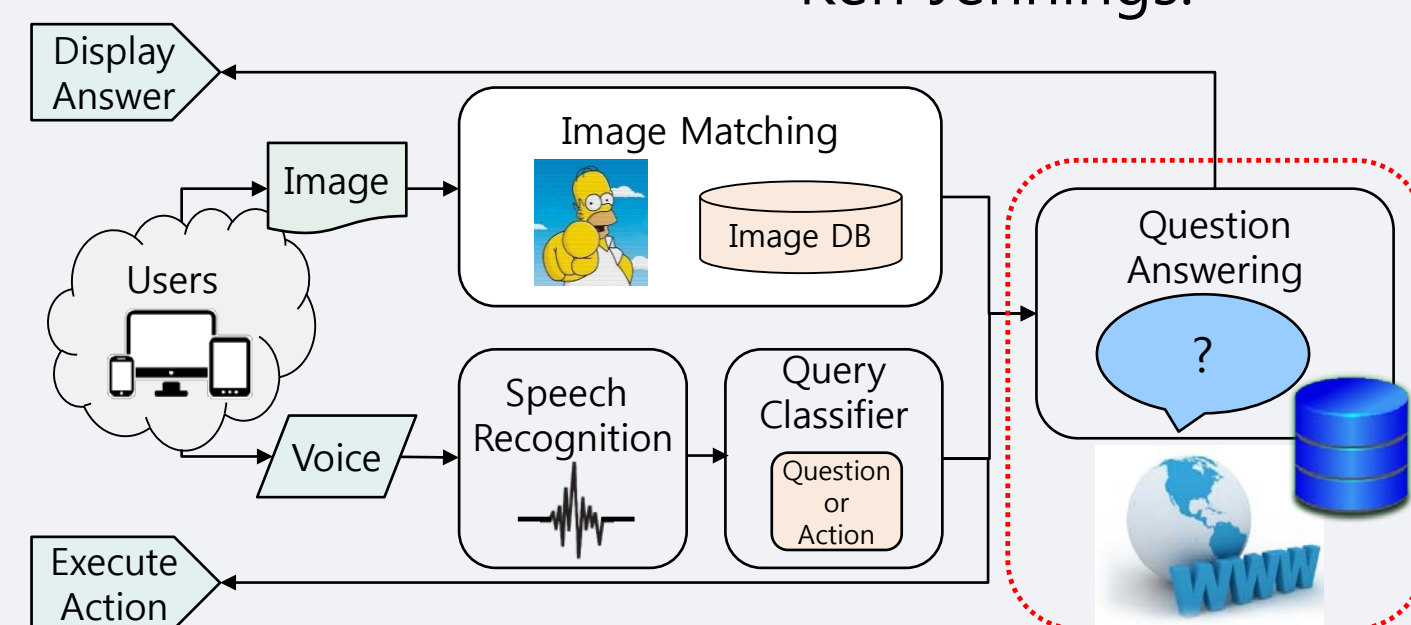


Question Answering Systems



Question answering system automatically answers questions posed by humans in a natural language.

For example, **IBM Watson** received the first place prize of \$1 million by competing on the quiz show Jeopardy! against former winners Brad Rutter and Ken Jennings.



The question answering system based on the **deep learning** is an interesting research topic.

Distributed and Parallel Algorithms



MapReduce is a framework for processing parallelizable problems across huge datasets using a large number of computers. It is useful in a wide range of applications, including distributed pattern-based searching, distributed sorting, document clustering, machine learning, and so on.

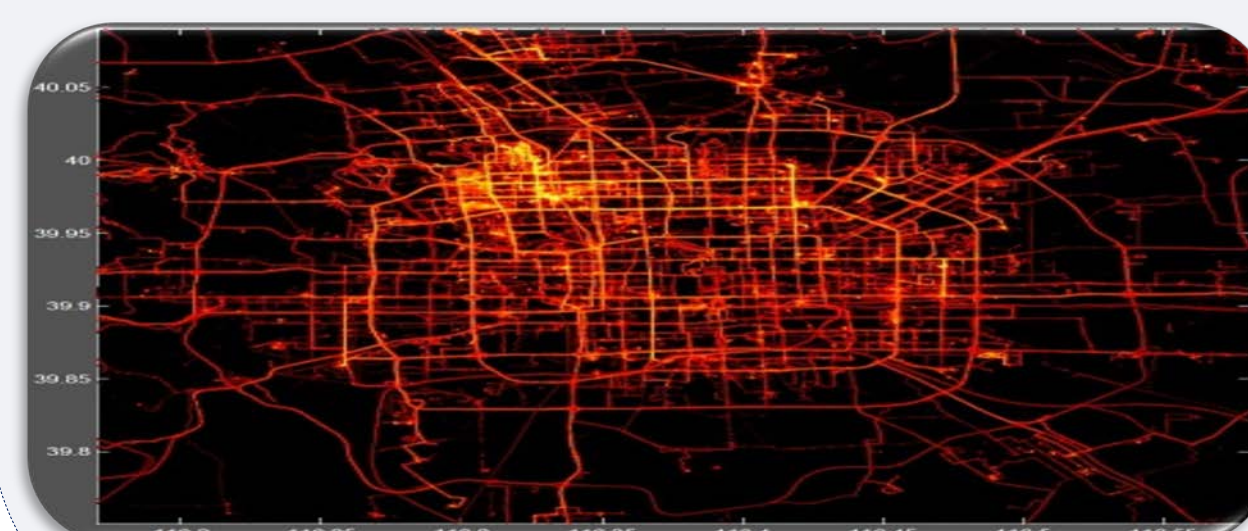


Big data analysis technology has an important role to find patterns of nature. It allows us to understand the past and predict the future.

Pattern Mining



In our ordinary life, there are several patterns around us. If we can find such **patterns**, then we could utilize them. For example, **crime patterns** detected from data enables us to predict crimes and aid the policemen.



Traffic pattern mining can be used in urban planning and many other applications