

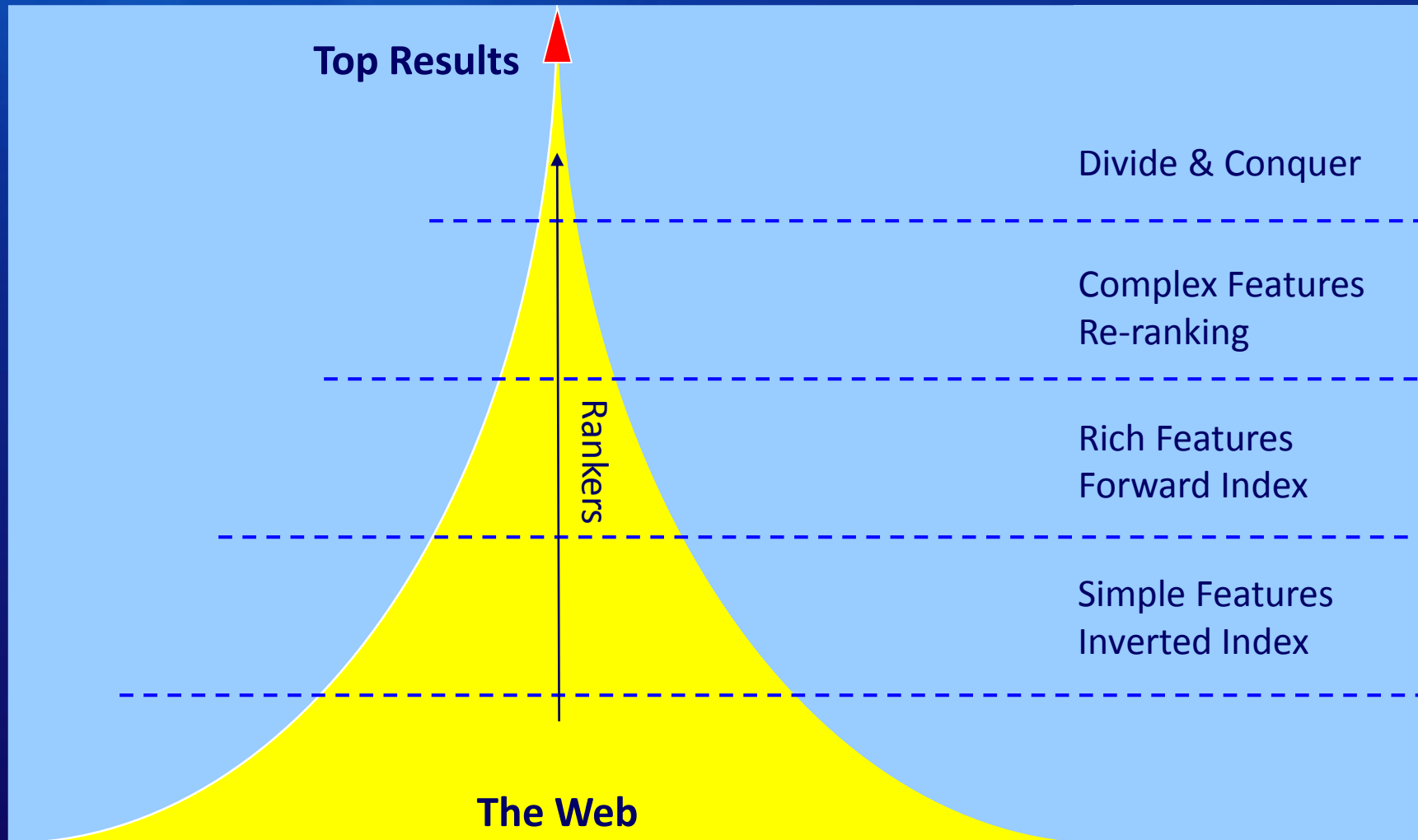
Advancing Web Search with New Capabilities from Knowledge Mining, Deep Learning, and Crowdsourcing

马维英 博士

微软亚洲研究院 常务副院长

Search Engine

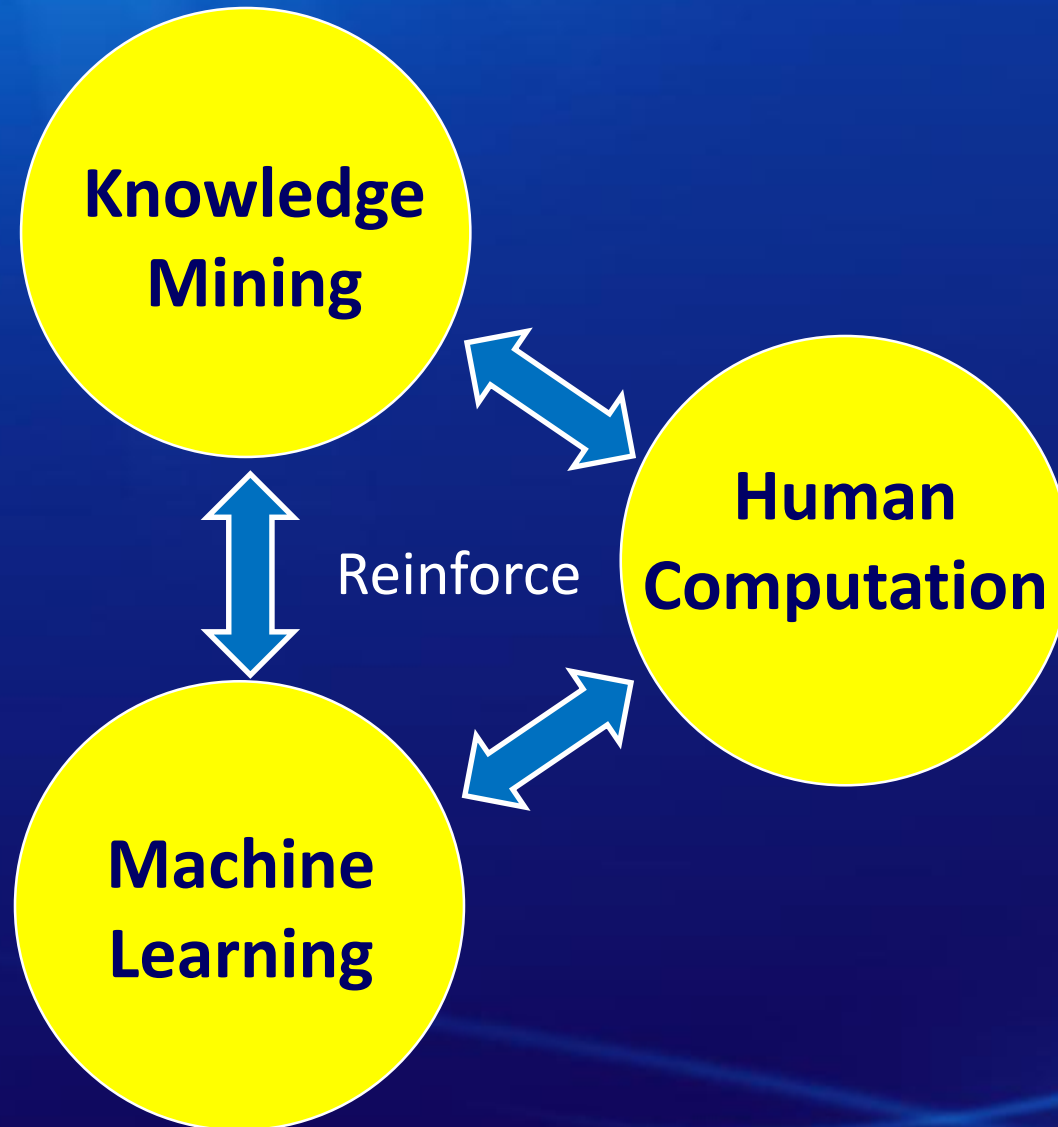
Query → Indexing → Documents → **Ranking**



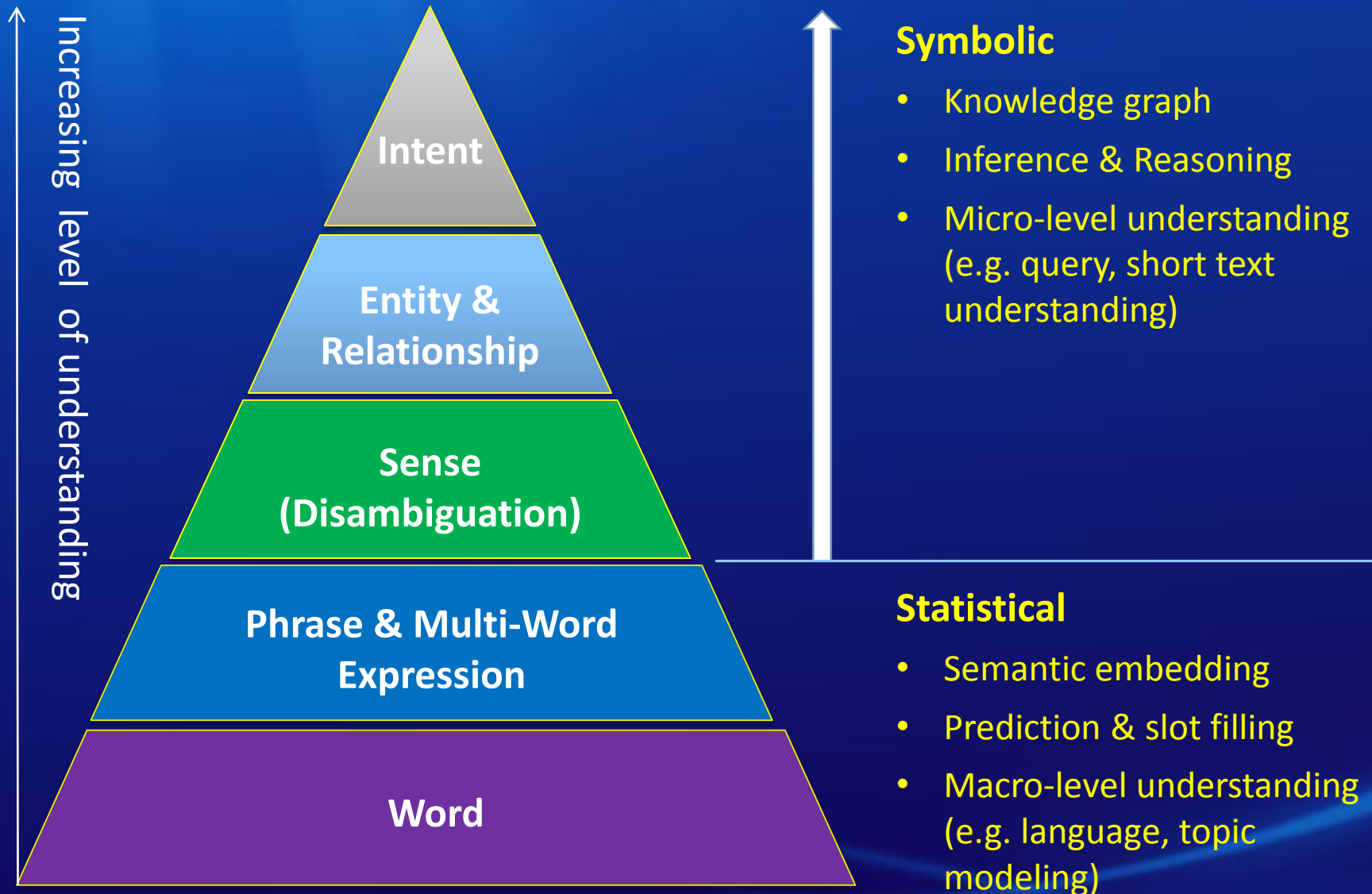
New Capabilities

- **Deep learning** (bottom-up)
 - Learn representation (feature) from data
 - Big data + Big infrastructure -> Big model + Big learning
- **Knowledge mining** (top-down)
 - Entities and relations, alternative expressions, etc.
 - Inference and reasoning
- **Crowdsourcing** (human computation in the loop)
 - Assist automatic algorithm to achieve higher accuracy
 - Treat human computer interaction as new way to collect more signals and labelled data for supervised learning

A “Moore’s Law” Phenomenon



Understand Document and Query



Knowledge Mining

Goal: Understand & serve the world through knowledge discovery and data mining

Data in the world

Nicolas and Carla Bruni-Sarkozy
Welcome a Daughter
By LIZ RAFTERY
10/19/2011 at 04:00 PM EDT

f Share Tweet f Like 726 +1 1 0 Comments



French First Lady Carla Bruni-Sarkozy gave birth to a daughter – her first child with French President Nicolas Sarkozy – on Wednesday in Paris, *Le Figaro* reports.

Bruni-Sarkozy, 43, who confirmed her pregnancy in July, delivered around 8 p.m. at the Clinique de la Muette after arriving earlier in the day.

The baby girl, whose name is unknown, is the first child born to a ruling French leader since Napoleon.

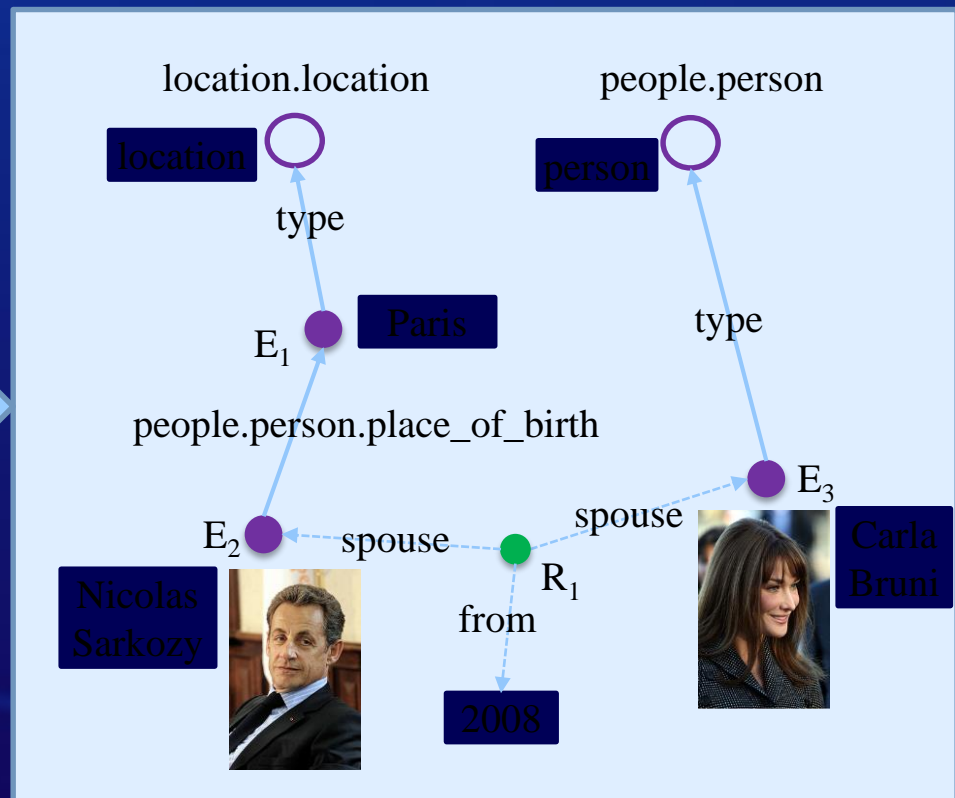
Prior to giving birth, the singer/actress said she will not be releasing photos of the baby, insisting that the French people would not be interested in them anyway.

Keep up with your favorite celebs in the pages of *PEOPLE* Magazine by [subscribing now](#).

RELATED: Carla Bruni-Sarkozy: France Is Not Interested in My Baby

Bruni-Sarkozy married the French President, 56, in February 2008. Sarkozy has three children from two previous marriages; Bruni-Sarkozy has one son from a previous relationship.

Entity Graph



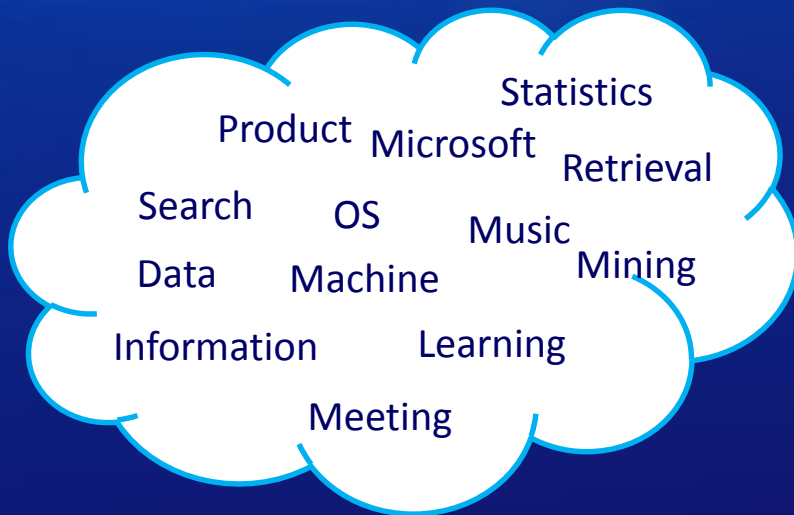
Knowledge Serving and the “Brain” in the Cloud



Collect ➡ Store ➡ Aggregate ➡ Analyze ➡ Organize ➡ Present ➡ Act

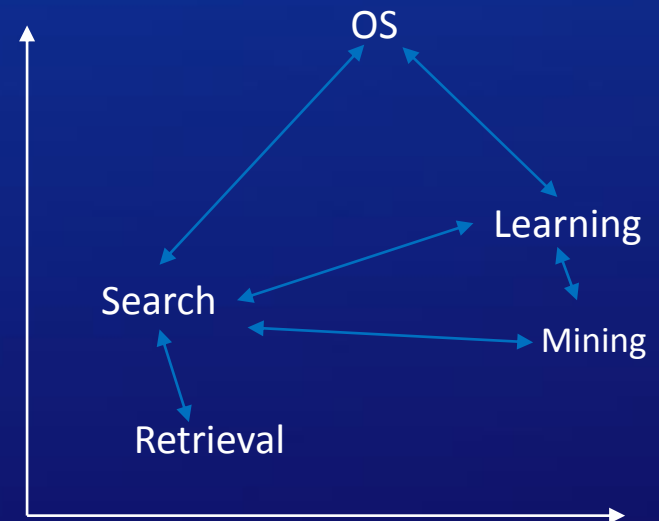
Word Embedding

Original Word Space



No meaningful relationship in original word space

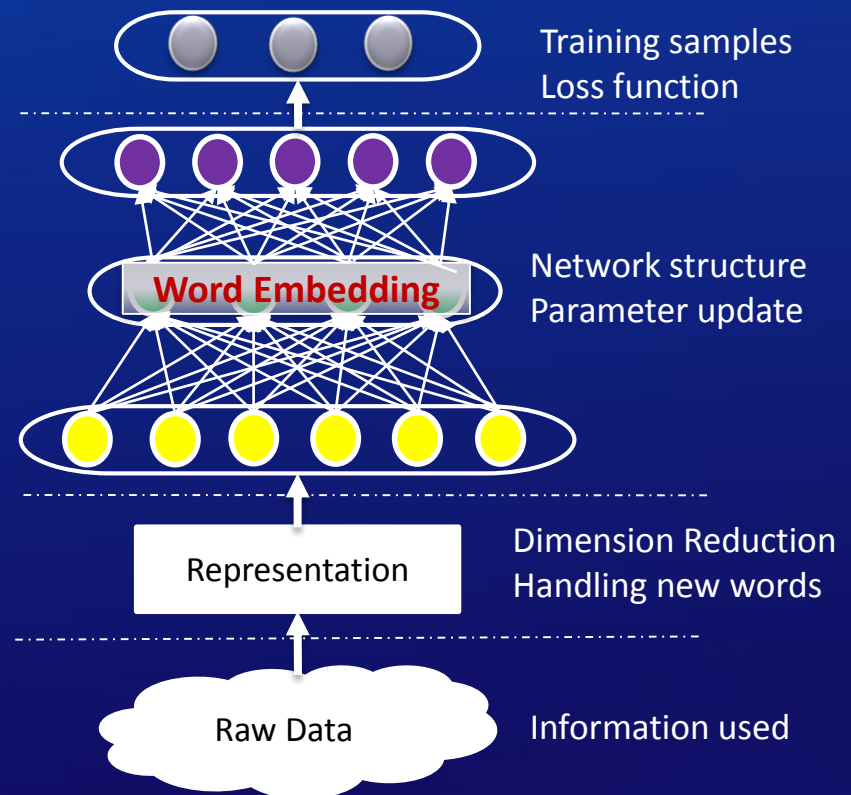
Embedding Space



Closer if semantically correlated in embedding space

Deep Learning for Word Embedding

4. Objective	$L(f(z), y)$
3. Model	$f(z)$
2. Representation	$z = \Phi(x)$
1. Data	x



Important Applications

- Query understanding
- Document understanding
- Information Extraction
- Question answering
- Machine translation
- Multimedia search
- Social media analytics
- Urban informatics

Summary

- Capabilities improved dramatically
 - Knowledge mining (top-down)
 - Deep learning (bottom-up)
 - Crowdsourcing (human computation in the loop)
- Take advantage of the new “Moore’s law” phenomenon
- Opportunities to advance search & NLP technologies