From Big Text To Big Knowledge!

Machine And Language Learning (MALL) Lab
SERC, Indian Institute of Science
It’s the IBM Watson!
Google Knowledge Graph!

How did Google know that Narendra Modi is the PM of India?
What do you think I meant in this tweet?

Dhawal Joharapurkar @dhawaljoh · 23s
I saw a penguin with binoculars!

Saw such a penguin?

Used binoculars to see the penguin?
How do humans process information?

Humans learn from experience

Background Knowledge!
Lots of it!
But, how does the computer understand?

Dhawal Joharapurkar @dhawaljoh · 23s
I saw a penguin with binoculars!

I used these!

Visit a doctor if you saw that!
Background world knowledge is key to Intelligent Decision Making

Where computers get a knowledge base (KB) of world facts?

- Manual entry doesn’t scale
- Automatic approach: read the web on a never-ending basis

A knowledge base (KB) is a technology used to store complex structured and unstructured information used by a computer system.
Challenges

“We’re still at 1 percent of where we should be.”

-- Larry Page (Google CEO) on Knowledge Graph [Jan 22, 2013]
Explosion of Unstructured Text!

300 million new websites added in 2011 alone (a 117% growth)

500 million Tweets per day (circa Oct 2012)
Time for one person to read all this data: 31 years
Automatically build, maintain, and make the KB available to intelligent applications (e.g., social media analysis, robotics, etc.)
• Persistent software individual (runs 24x7x365)
• Learns many functions / knowledge types
• Learns easier things first, then more difficult ones
• The more it learns, the more it can learn next
• Learns from experience and from advice
• Goal-directed KB expansion
• Continuous KB evaluation
• Temporal Micro Reading
• Representation Learning
• Knowledge on Demand
Unprecedented opportunity to bring world knowledge into AI systems -- focus of our research
Thank You!

Dhawal Joharapurkar

MALL Lab: www.talukdar.net/mall-lab.html