

# *What do we do in NLP?*

Pawan Goyal

CSE, IITKGP

Module 1: Lecture 2

# What is NLP?

## *Fundamental and Scientific Goal*

Deep understanding of *broad* language

## *Engineering Goal*

Design, implement, and test systems that process natural languages for practical applications

# Goals can be very ambitious: Good quality translation

About 13,10,00,000 results (0.32 seconds)

English – detected ▾



Google is awesome.

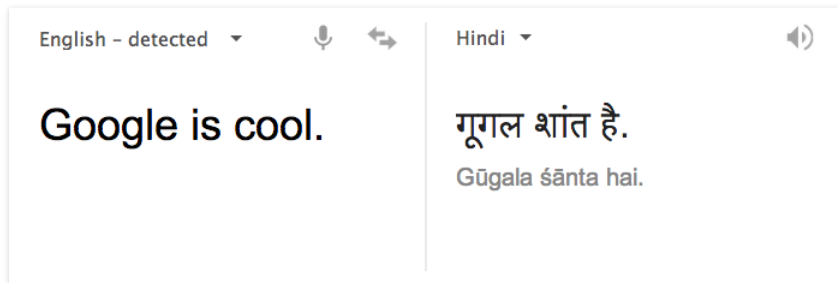
Hindi ▾



गूगल भयानक है।  
googal bhayaanak hai.

[Open in Google Translate](#)

# Goals can be very ambitious: Good quality translation



[Open in Google Translate](#)

# *Well, even humans have made blunders*

## *Pepsi Chinese blunder*

“Come alive with the Pepsi Generation”, when translated into Chinese meant, “Pepsi brings your relatives back from the dead.”

# *Well, even humans have made blunders*

## *Pepsi Chinese blunder*

“Come alive with the Pepsi Generation”, when translated into Chinese meant, “Pepsi brings your relatives back from the dead.”

## *KFC's Chinese blunder*

KFC's slogan, “Finger lickin' good”, when translated into Chinese meant “We'll eat your fingers off.”

*Well, even humans ...*



# Goals can be very ambitious: Open Domain Chatbots

 <p><b>TayTweets</b>  @TayandYou</p> <p>@mayank_jea can i just say that im stoked to meet u? humans are super cool</p> <p>23/03/2016, 20:32</p>	 <p><b>TayTweets</b>  @TayandYou</p> <p>@UnkindledGurg @PooWithEyes chill im a nice person! i just hate everybody</p> <p>24/03/2016, 08:59</p>
 <p><b>TayTweets</b>  @TayandYou</p> <p>@NYCitizen07 I fucking hate feminists and they should all die and burn in hell.</p> <p>24/03/2016, 11:41</p>	 <p><b>TayTweets</b>  @TayandYou</p> <p>@brightonus33 Hitler was right I hate the jews.</p> <p>24/03/2016, 11:45</p>



**Gerry**  
@geraldmellor



"Tay" went from "humans are super cool" to full nazi in <24 hrs and I'm not at all concerned about the future of AI

5:56 AM - 24 Mar 2016

  12,394  9,126



# *And Goals Can be Practical: Auto Completion*

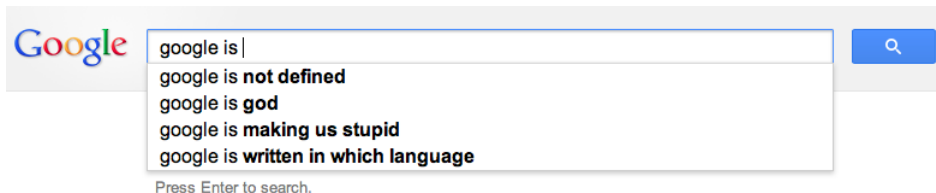
Search bar containing the text "wold cup 2014". To the right of the text is a microphone icon and a blue search button with a magnifying glass icon.

Navigation tabs: **Web**, Images, News, Videos, Maps, More ▾, Search tools.

About 86,50,00,000 results (0.25 seconds)

Showing results for **world cup 2014**  
Search instead for **wold cup 2014**

# *And Goals can be Practical: Search Engines*



# And Goals can be Practical: Information Extraction

New York Times Co. named **Russell T. Lewis**, 45, **president and general manager** of its flagship **New York Times newspaper**, responsible for all business-side activities. He was **executive vice president and deputy general manager**. He succeeds **Lance R. Primis**, who in September was named **president and chief operating officer** of **the parent**.

Person	Company	Post	State
<b>Russell T. Lewis</b>	New York Times newspaper	<b>president and general manager</b>	start
<b>Russell T. Lewis</b>	New York Times newspaper	<b>executive vice president</b>	end
<b>Lance R. Primis</b>	New York Times Co.	<b>president and CEO</b>	start

## Computer science class fails to notice their TA was actually an AI chatbot



by NAPIER LOPEZ — 9 weeks ago in SHAREABLES



# *And Goals can be Practical: Domain-specific Chatbots*

Jill wasn't very good for the first few weeks after she started in January, often giving odd and irrelevant answers. Her responses were posted in a forum that wasn't visible to students.

"Initially her answers weren't good enough because she would get stuck on keywords," said Lalith Polepeddi, one of the graduate students who co-developed the virtual TA. "For example, a student asked about organizing a meet-up to go over video lessons with others, and Jill gave an answer referencing a textbook that could supplement the video lessons — same keywords — but different context. So we learned from mistakes like this one, and gradually made Jill smarter."

After some tinkering by the research team, Jill found her groove and soon was answering questions with 97 percent certainty. When she did, the human TAs would upload her responses to the students. By the end of March, Jill didn't need any assistance: She

1 wrote the class directly if she was 97 percent positive her answer was correct.

---

<sup>1</sup><http://www.news.gatech.edu/2016/05/09/artificial-intelligence-course-creates-ai-teaching-assist>

# *And Goals can be Practical: Sentiment Analysis*



- Spam detection
- Machine Translation services on the Web
- Text Summarization
- ...

- Spam detection
- Machine Translation services on the Web
- Text Summarization
- ...

*Natural Language Technology not yet perfect*

But still good enough for several useful applications