NLP

Introduction to NLP

NLP Tasks (cont'd)

Question Answering

- **"The antagonist of Stevenson's Treasure Island." (Who is Long John Silver?)**
- http://blog.reddit.com/2011/02/ibm-watson-research-team-answers-your.html
- "Watson is powered by 10 racks of IBM Power 750 servers running Linux, and uses 15 terabytes of RAM, 2,880 processor cores and is capable of operating at 80 teraflops. Watson was written in mostly Java but also significant chunks of code are written C++ and Prolog, all components are deployed and integrated using UIMA."

Jeopardy Questions

- From the competition between the IBM Watson system and two human champions (Ken Jennings and Brad Rutter)
- Sample questions:
  - On December 8, 2008 this national newspaper raised its newsstand price by 25 cents to $1 : *USA Today*
  - In 2010 this former first lady published the memoir "Spoken From the Heart" : *Laura Bush**
  - This person is appointed by a testator to carry out the directions & requests in his will : *Executor**
  - Familiarity is said to breed this, from the Latin for "Despise" : *Contempt**
  - As of 2010, Croatia & Macedonia are candidates but this is the only former Yugoslav republic in the EU : *Slovenia*
  - The ancient "Lion of Nimrud" went missing from this city's national museum in 2003 (along with a lot of other stuff) : *Baghdad*
  - It's just a bloody nose! You don't have this hereditary disorder once endemic to European royalty : *Haemophilia*
  - It's Michelangelo's fresco on the wall of the Sistine Chapel, Depicting the saved and the damned : *The Last Judgement*
  - She "Died in the church and was buried along with her name. Nobody came" : *Eleanor Rigby*
  - It's a 4-letter term for a summit; the first 3 letters mean a type of simian : *Apex*
  - A camel is a horse designed by this : *Committee*
- Watson's answers: 66 correct and 9 incorrect (e.g., the one in the category "US Cities" about a city with two airports named after a World War II hero and a World War II battle)
- Watson's two day winning streak was $77,147. Ken Jennings ended with $24,000 and Brad Rutter with $21,600. Watson donated $500,000 to both World Vision and World Community Grid charities from the $1,000,000 prize.
- http://www.quora.com/What-questions-were-asked-in-the-Jeopardy-episode-involving-Watson

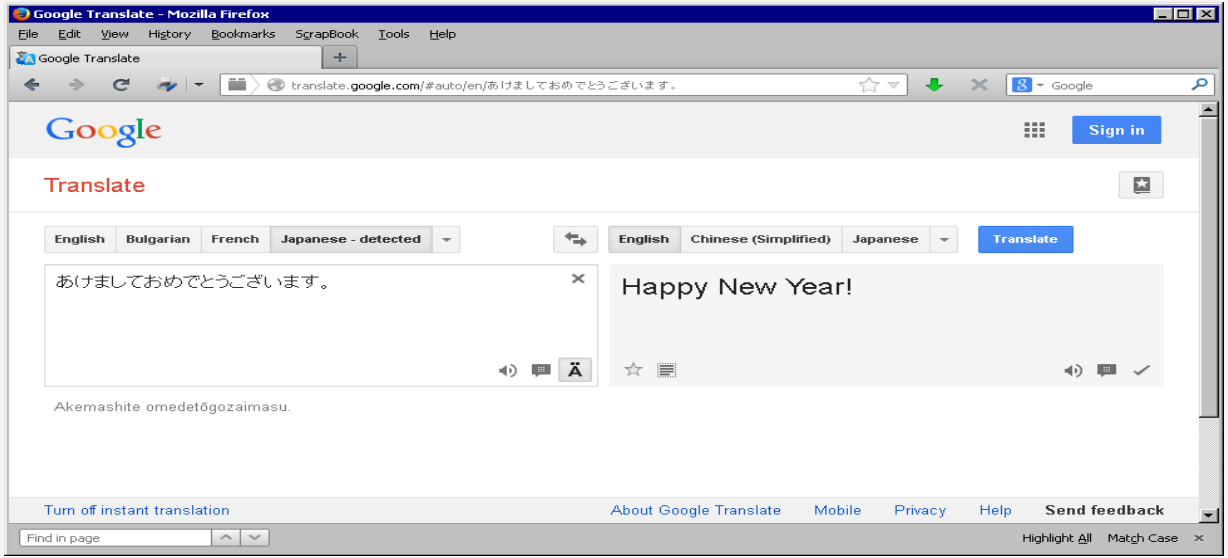
Sentiment Analysis

- "I like the camera because I can edit images so easily, exactly as I do my iPad. I have found that its difficult to frame a picture when there isn't a zoom function as with the iPad. With this camera I can adjust my images by cropping as I did with my iPad but better yet, this camera has a built in zoom. A stretch or pinch of the fingers bring in the subject closer or back out again. With this iPhone I can also, as I dido with my iPad, enhance, crop, rotate, red eye reduce, and set a range of tints. I am also quite impressed with the quality of the images. Pretty darn good especially better than I expected for low light situations where I can use the built-in flash! Quite frankly I was quite surprised with these built in features. I also hope too experiment with and learn what HDR photography is. It's built into this iPhone and can be activated by a the touch of an icon. "

- http://www.epinions.com/review/apple_iphone_5c_latest_model_16gb_graphite_unlocked_smartphone/content_640679317124

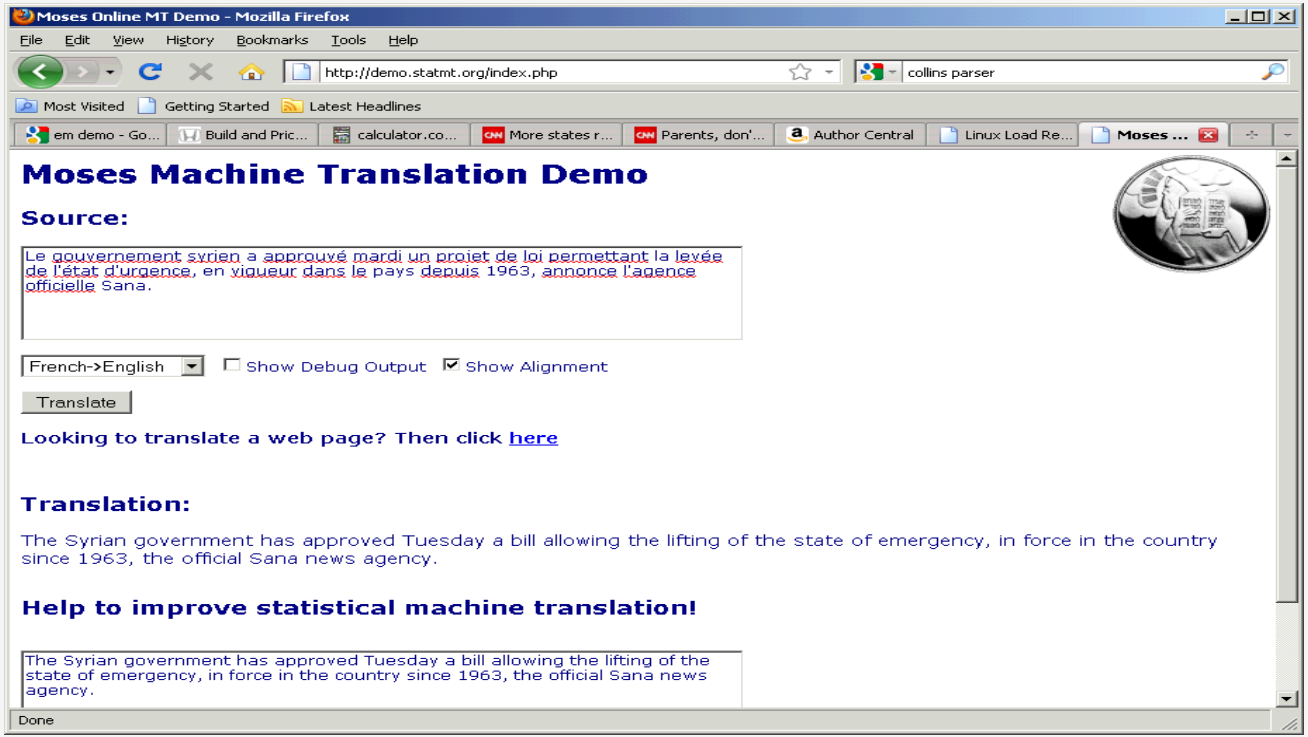
Machine Translation

- あけましておめでとうございます。
- Happy New Year!



Machine Translation

- Moses
- www.statmt.org



Elephants are social animals. They live with their families, give hugs and call each other by using their trunks as trumpets. They also might know how to help each other.

In a recent elephant study by researchers from the United States and Thailand, pairs of giant animals learned to work together to get some ears of corn. Other animals, especially some primates, are already known to work together to complete tasks, but now elephants have joined the club. Perhaps the finding is not too surprising: Scientists suspect that elephants, with their big brains and survival savvy, may be among the smartest animals on the planet.

Joshua Plotnik, who worked on the study, told Science News that the animals didn't just learn a trick. Instead, the ways the elephants behaved show that they understand how working together brings benefits to everyone involved. Plotnik is a comparative psychologist now at the University of Cambridge in England. Psychology is the study of behaviors and mental processes, and comparative psychologists study how animals other than humans behave.

Les éléphants sont des animaux sociaux. Ils vivent avec leur famille, faire des câlins et appeler les uns les autres en utilisant leurs troncs trompettes. Ils pourraient également savoir comment aider les uns les autres.

Dans une étude récente d'éléphants par des chercheurs des États-Unis et la Thaïlande, des paires d'animaux géants ont appris à travailler ensemble pour obtenir des épis de maïs. D'autres animaux, en particulier des primates, sont déjà connus pour travailler ensemble pour accomplir des tâches, mais maintenant, les éléphants ont rejoint le club. Peut-être le résultat n'est pas trop surprenant: Les scientifiques soupçonnent que les éléphants, avec leurs gros cerveaux et de bon sens de survie, peut-être parmi les plus intelligents des animaux sur la planète.

Joshua Plotnick, qui a travaillé sur l'étude, dit Nouvelles de la Science que les animaux n'ont pas seulement appris un truc. Au lieu de cela, les moyens les éléphants se comportent montrent qu'ils comprennent comment travailler ensemble apporte des avantages à toutes les personnes impliquées. Plotnik est un psychologue comparative maintenant à l'Université de Cambridge en Angleterre. **La psychologie est l'étude des comportements et des processus mentaux**, et étude comparative des psychologues comment les animaux autres que les humains se comportent.

https://student.societyforscience.org/article/theres-no-i-elephant

Elephants are social animals. They live with their families, give hugs and call each other by using their trunks as trumpets. They also might know how to help each other.

In a recent elephant study by researchers from the United States and Thailand, pairs of giant animals learned to work together to get some ears of corn. Other animals, especially some primates, are already known to work together to complete tasks, but now elephants have joined the club. Perhaps the finding is not too surprising: Scientists suspect that elephants, with their big brains and survival savvy, may be among the smartest animals on the planet.

Joshua Plotnik, who worked on the study, told Science News that the animals didn't just learn a trick. Instead, the ways the elephants behaved show that they understand how working together brings benefits to everyone involved. Plotnik is a comparative psychologist now at the University of Cambridge in England. Psychology is the study of behaviors and mental processes, and comparative psychologists study how animals other than humans behave.

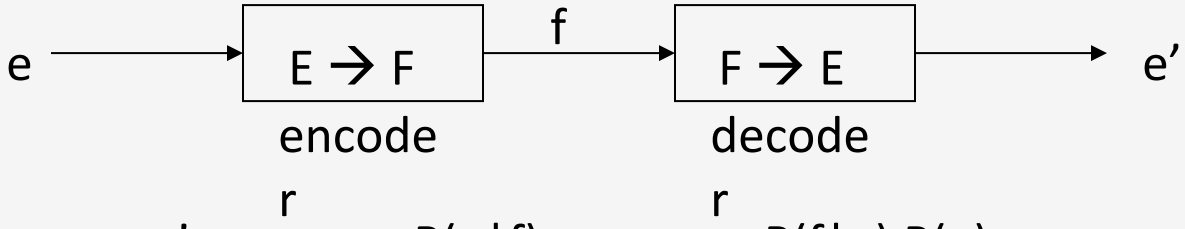
Les éléphants sont des animaux sociaux. Ils **vivent** avec leur famille, **faire** des câlins et **appeler** les uns les autres en utilisant leurs troncs trompettes. Ils pourraient également savoir comment aider les uns les autres.

Dans une étude récente d'éléphants par des chercheurs des États-Unis et la Thaïlande, des paires d'animaux géants ont appris à travailler ensemble pour obtenir des épis de maïs. D'autres animaux, en particulier des primates, sont déjà connus pour travailler ensemble pour accomplir des tâches, mais maintenant, les éléphants ont rejoint le club. Peut-être le résultat n'est pas trop surprenant: Les scientifiques soupçonnent que **les éléphants**, avec leurs gros cerveaux et de bon sens de survie, **peut-être** parmi les plus intelligents des animaux sur la planète.

Joshua Plotnick, qui a travaillé sur l'étude, **dit Nouvelles de la Science** que les animaux n'ont pas seulement appris un truc. Au lieu de cela, les moyens les éléphants se comportent montrent qu'ils comprennent comment travailler ensemble apporte des avantages à toutes **les personnes** impliquées. Plotnik est un psychologue **comparative** maintenant à l'Université de Cambridge en Angleterre. La psychologie est l'étude des comportements et des processus mentaux, et **étude comparative des psychologues** comment les animaux autres que les humains se comportent.

Machine Translation

- Noisy channel model ("Chinese Whispers")



Machine Translation

- IBM Method

Text Summarization

Health Benefits

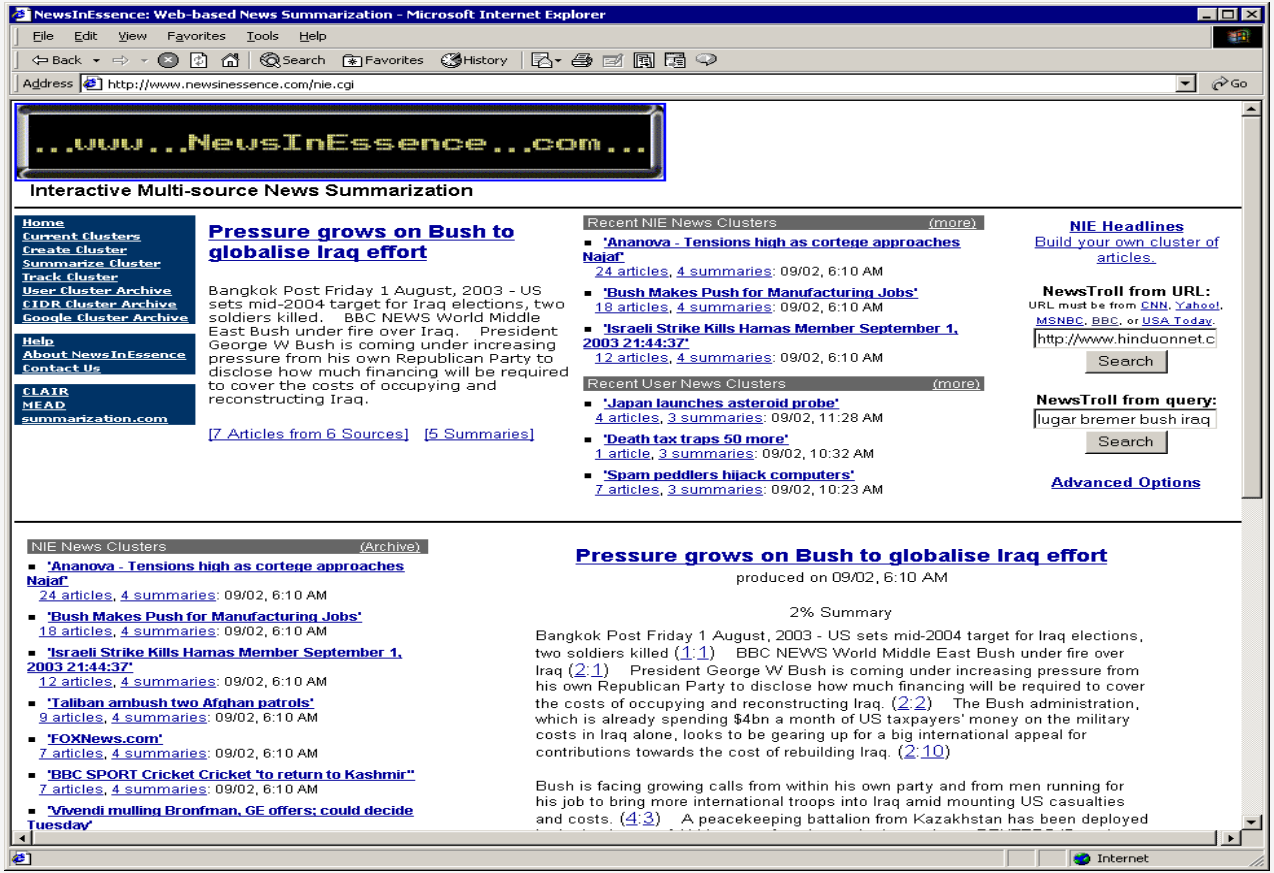
- Eating a diet rich in vegetables and fruits as part of an overall healthy diet may reduce risk for heart disease, including heart attack and stroke.
- Eating a diet rich in some vegetables and fruits as part of an overall healthy diet may protect against certain types of cancers.
- Diets rich in foods containing fiber, such as some vegetables and fruits, may reduce the risk of heart disease, obesity, and type 2 diabetes.
- Eating vegetables and fruits rich in potassium as part of an overall healthy diet may lower blood pressure, and may also reduce the risk of developing kidney stones and help to decrease bone loss.
- Eating foods such as vegetables that are lower in calories per cup instead of some other higher-calorie food may be useful in helping to lower calorie intake.

Nutrients

- Most vegetables are naturally low in fat and calories. None have cholesterol. (Sauces or seasonings may add fat, calories, or cholesterol.)
- Vegetables are important sources of many nutrients, including potassium, dietary fiber, folate (folic acid), vitamin A, and vitamin C.
- Diets rich in potassium may help to maintain healthy blood pressure. Vegetable sources of potassium include sweet potatoes, white potatoes, white beans, tomato products (paste, sauce, and juice), beet greens, soybeans, lima beans, spinach, lentils, and kidney beans.
- Dietary fiber from vegetables, as part of an overall healthy diet, helps reduce blood cholesterol levels and may lower risk of heart disease. Fiber is important for proper bowel function. It helps reduce constipation and diverticulosis. Fiber-containing foods such as vegetables help provide a feeling of fullness with fewer calories.
- Folate (folic acid) helps the body form red blood cells. Women of childbearing age who may become pregnant should consume adequate folate from foods, and in addition 400 mcg of synthetic folic acid from fortified foods or supplements. This reduces the risk of neural tube defects, spina bifida, and anencephaly during fetal development.

Summary

Eating vegetables is healthy.


NewsInEssence: Web-based News Summarization - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://www.newsinessence.com/nie.cgi
...www...NewsInEssence...com...
Interactive Multi-source News Summarization
Home
Current Clusters
Create Cluster
Summarize Cluster
Track Cluster
User Cluster Archive
CIDR Cluster Archive
Google Cluster Archive
Help
About NewsInEssence
Contact Us
CLAIR
MEAD
summarization.com
Pressure grows on Bush to globalise Iraq effort
Bangkok Post Friday 1 August, 2003 - US sets mid-2004 target for Iraq elections, two soldiers killed. BBC NEWS World Middle East Bush under fire over Iraq. President George W Bush is coming under increasing pressure from his own Republican Party to disclose how much financing will be required to cover the costs of occupying and reconstructing Iraq.
[7 Articles from 6 Sources] [5 Summaries]
Recent NIE News Clusters (more)
'Ananova - Tensions high as cortege approaches Najaf'
24 articles, 4 summaries: 09/02, 6:10 AM
'Bush Makes Push for Manufacturing Jobs'
18 articles, 4 summaries: 09/02, 6:10 AM
'Israeli Strike Kills Hamas Member September 1, 2003 21:44:37'
12 articles, 4 summaries: 09/02, 6:10 AM
Recent User News Clusters (more)
'Japan launches asteroid probe'
4 articles, 3 summaries: 09/02, 11:28 AM
'Death tax traps 50 more'
1 article, 3 summaries: 09/02, 10:32 AM
'Spam peddlers hijack computers'
7 articles, 3 summaries: 09/02, 10:23 AM
NIE Headlines
Build your own cluster of articles.
NewsTroll from URL:
URL must be from CNN, Yahoo!, MSNBC, BBC, or USA Today.
http://www.hinduonnet.c
Search
NewsTroll from query:
lugar bremer bush iraq
Search
Advanced Options
NIE News Clusters (Archive)
'Ananova - Tensions high as cortege approaches Najaf'
24 articles, 4 summaries: 09/02, 6:10 AM
'Bush Makes Push for Manufacturing Jobs'
18 articles, 4 summaries: 09/02, 6:10 AM
'Israeli Strike Kills Hamas Member September 1, 2003 21:44:37'
12 articles, 4 summaries: 09/02, 6:10 AM
'Taliban ambush two Afghan patrols'
9 articles, 4 summaries: 09/02, 6:10 AM
'FOXNews.com'
7 articles, 4 summaries: 09/02, 6:10 AM
'BBC SPORT Cricket Cricket 'to return to Kashmir''
7 articles, 4 summaries: 09/02, 6:10 AM
'Vivendi mulling Bronfman, GE offers; could decide Tuesday'
Pressure grows on Bush to globalise Iraq effort
produced on 09/02, 6:10 AM
2% Summary
Bangkok Post Friday 1 August, 2003 - US sets mid-2004 target for Iraq elections, two soldiers killed (1:1) BBC NEWS World Middle East Bush under fire over Iraq (2:1) President George W Bush is coming under increasing pressure from his own Republican Party to disclose how much financing will be required to cover the costs of occupying and reconstructing Iraq. (2:2) The Bush administration, which is already spending $4bn a month of US taxpayers' money on the military costs in Iraq alone, looks to be gearing up for a big international appeal for contributions towards the cost of rebuilding Iraq. (2:10)
Bush is facing growing calls from within his own party and from men running for his job to bring more international troops into Iraq amid mounting US casualties and costs. (4:3) A peacekeeping battalion from Kazakhstan has been deployed
Internet

Text to Speech

http://www2.research.att.com/~ttsweb/tts/demo.php

Text to Speech

- www.ivona.com

Entailment and Paraphrasing

ID	TEXT	HYPOTHESIS	TASK	VALUE
1586	*The Republic of Yemen is an Arab, Islamic and independent sovereign state whose integrity is inviolable, and no part of which may be ceded.*	*The national language of Yemen is Arabic.*	QA	True
1076	*Most Americans are familiar with the Food Guide Pyramid– but a lot of people don't understand how to use it and the government claims that the proof is that two out of three Americans are fat.*	*Two out of three Americans are fat.*	RC	True
1667	*Regan attended a ceremony in Washington to commemorate the landings in Normandy.*	*Washington is located in Normandy.*	IE	False
2016	*Google files for its long awaited IPO.*	*Google goes public.*	IR	True
2097	*The economy created 228,000 new jobs after a disappointing 112,000 in June.*	*The economy created 228,000 jobs after disappointing the 112,000 of June.*	MT	False
893	*The first settlements on the site of Jakarta were established at the mouth of the Ciliwung, perhaps as early as the 5th century AD.*	*The first settlements on the site of Jakarta were established as early as the 5th century AD.*	CD	True
1960	*Bush returned to the White House late Saturday while his running mate was off campaigning in the West.*	*Bush left the White House.*	PP	False
586	*The two suspects belong to the 30th Street gang, which became embroiled in one of the most notorious recent crimes in Mexico: a shootout at the Guadalajara airport in May, 1993, that killed Cardinal Juan Jesus Posadas Ocampo and six others.*	*Cardinal Juan Jesus Posadas Ocampo died in 1993.*	QA	True

Table 1. Examples of Text-Hypothesis pairs

Ido Dagan, Oren Glickman and Bernardo Magnini. The PASCAL Recognising Textual Entailment Challenge

Discourse Analysis

- Anaphoric relations:

1. Mary helped Peter get out of the car. He thanked her.
2. Mary helped the other passenger out of the car.
The man had asked her for help because of his foot injury.

Tom appeared on the sidewalk with a bucket of whitewash and a long-handled brush. He surveyed the fence, and all gladness left him and a deep melancholy settled down upon his spirit. (Tom Sawyer)

Dialogue Systems

- I would like to make a reservation at Sorrento.
- For when?
- 8 pm Friday night.
- We only have availability for 7 pm and 10 pm.
- Sorry, these don't work for me.

Other Applications

- Spelling Correction
- Web search
- Natural language interfaces to databases
- Parsing job postings
- Summarizing medical records
- Information extraction for databases
- Social network extraction from text
- Alignment of text w/ other signal (time series)
- Essay grading
- Generating weather reports, sports reports, and news stories

NLP