Fundamentals of AI: COMP3000
Natural Language Processing
Wentworth Institute of Technology

Natural Language Processing (NLP): Definition

Natural language processing is the ability to take a body of text and **extract meaning** from it using a computer.

Understanding Language

Video 4:22 minutes

Natural Language Challenge: Human vs Computer

HUMAN

"I drove my friend Mary to the park in my Tesla while listening to music on my iPhone."

Humans understand that Mary is a friend and that a Tesla is likely a car. Additionally, after many years of popularity and cultural references, we all know that an iPhone is a smartphone.

MACHINE

Structured data

```
<friend>Mary</friend>
<car>Tesla</car>
<phone>iPhone</phone>
```

None of this is understood by a computer without assistance.

NLP Components

Let's analyze the phrase …

"I drove my friend Mary to the park in my Tesla while listening to music on my iPhone"

Entities
The people, places, organizations, and things in your text.
Example: friend, car, and phone

Semantic Roles
Subjects, actions, and objects in the text.
"IBM bought a company." The subject is "IBM," the action is "bought," and the object is "company."

Relations
How entities are related.
A "createdBy" relation might connect the entities "iPhone" and "Apple."

Categories
Describing what a piece of content is about at a high level.
Categories could be sports, finance, travel, computing, and so on.

Concepts
Extracting reference to topics that do not explicitly appear in the text.
An article about Tesla may refer to concepts "electric cars" or "Elon Musk," even if those terms are not explicitly mentioned.

Emotion
Understanding the emotion or tone conveyed.
Is the content conveying anger, disgust, fear, joy, or sadness?

Keywords
Identify the important and relevant keywords in your content.

Sentiment
Is the feeling/attitude positive, neutral, or negative?
The level of positive or negative sentiment can be scored.

Natural Language Processing

The semantic behind the syntactic

Syntactic messages

Subject-verb-object

Semantic messages

Agent and Patient

What is the message sentiment?

SPECIAL COMMEMORATIVE SECTION

"All the News That's Fit to Print"

LATE CITY EDITION

VOL. CXVIII...No. 40,321 NEW YORK, MONDAY, JULY 21, 1969 10 CENTS

MEN WALK ON MOON

ASTRONAUTS LAND ON PLAIN; COLLECT ROCKS, PLANT FLAG

Voice From Moon: 'Eagle Has Landed'

A Powdery Surface Is Closely Explored

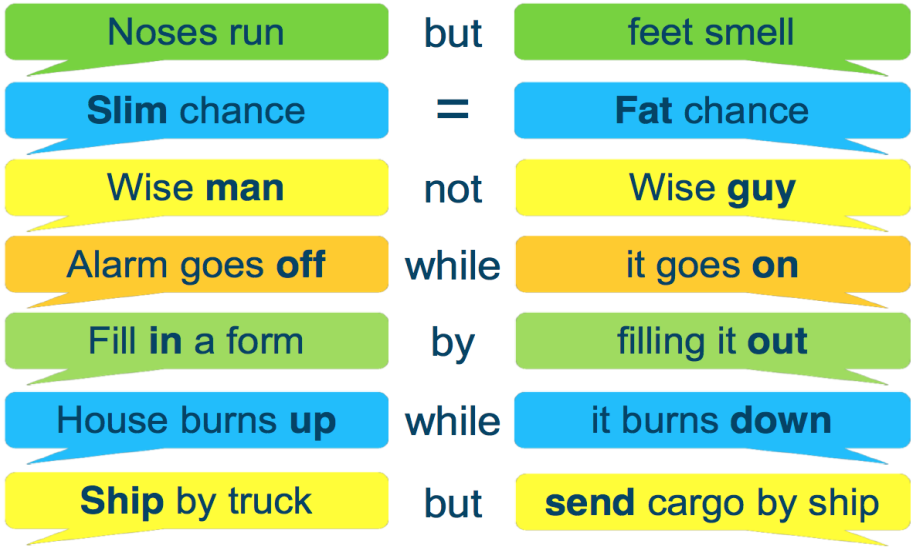
Tentative Steps Test Soil

Neil A. Armstrong moves away from the leg of the landing craft after taking the first step on the surface of the moon

* Search for 'moving' at http://www.shoecomics.com/

Natural language processing: a classification problem

- Difficulty of language: Subtleties, idiosyncrasies, idioms, ambiguities, nuances and gaps



- It is highly contextual, imprecise and has gaps (context known outside the conversation)

Enterprise Applications of NLP

- Social Media Analysis
- Customer Support
- Business Intelligence
- Content Marketing and Recommendation
- Additional Topics

Social Media Analysis

How can we extract valuable insights from social media posts?

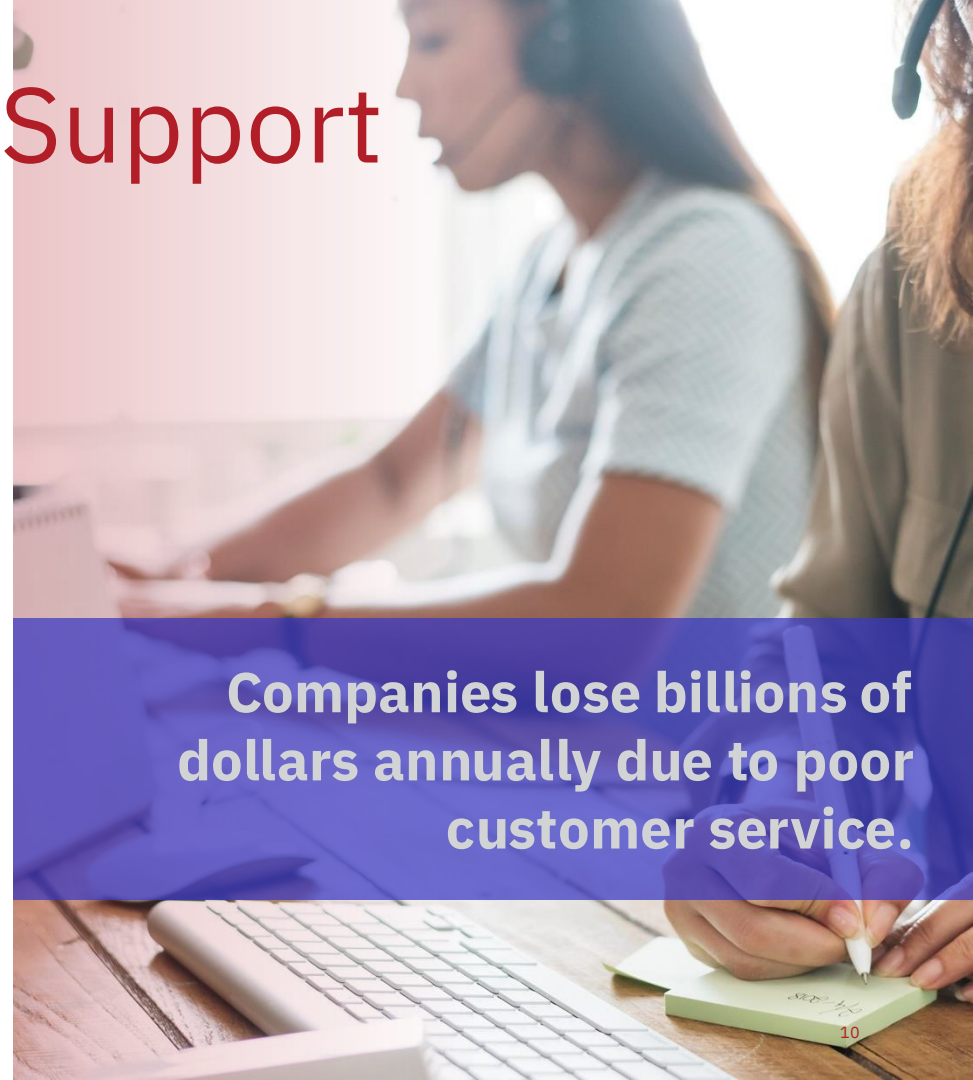
What are the relevant trending topics and hashtags for a business?

NLP can deliver this information and more.

Customer Support

Each support ticket can be analyzed to obtain its sentiment, keywords, and a categorization.

This process can be used to route the customer to the correct representative and in some cases to **automatically respond to the request**.



Business Intelligence

How does a company efficiently access unstructured data?

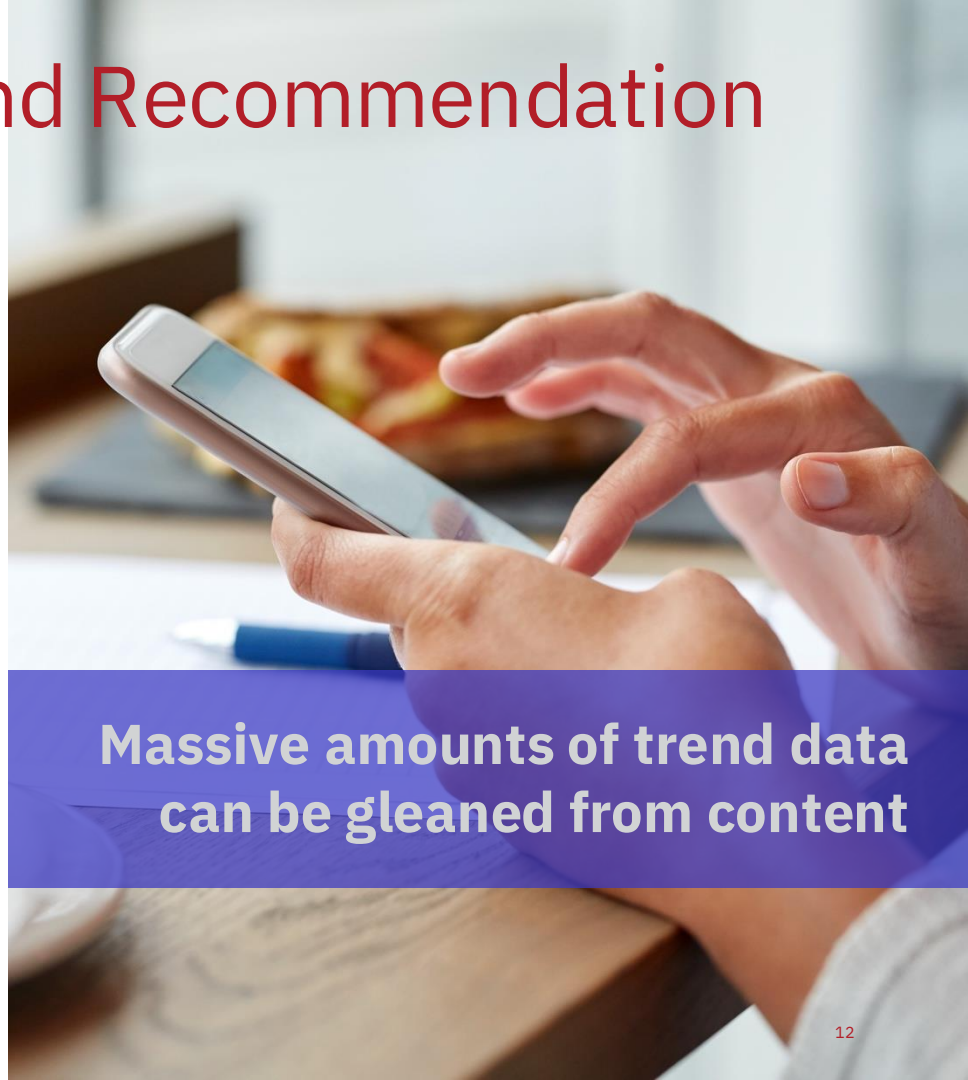
How can this data be queried on an ad hoc basis without the need for developers to write complex queries?

NLP allows all users to ask questions of the data without developer assistance.

Content Marketing and Recommendation

Companies also want better ways to recommend more relevant content to their readers.

NLP enables companies publishing content to understand what to write about as well as produce more interesting and relevant topics to readers



How to use NLP Today?

Open source software projects

The more popular include:

- **Apache NLP**
- **Stanford CoreNLP**
- **NLTK for Python**
- **SyntaxNet**

Software offered as a service

The more popular include:

- **Watson's Natural Language Understanding**
- **Azure Text Analytics**
- **Amazon's Lex**

NLP as core business*

The more popular include:

- 35% of Amazon Purchases
- 75% of Netflix Viewings

*Recommendation engine algorithms

What Is a **Virtual Agent or Chatbot?**

A chatbot is a way to expose a business's service or data via a natural language interface.

- **A virtual customer support agent can reduce headcount and exponentially scale real-time customer support capabilities.**
- **A conversational commerce chatbot gives business a whole new channel to engage with customers via messaging platforms.**

Factors in the Rise of Chatbots

Natural Language Processing in the Cloud

The availability of NLP capabilities in the cloud has been the most potent force behind the rise of chatbots.

Text classifiers and entity extractors power core functionalities inside a chatbot.

Proliferation of Messaging Platforms

Companies want to reach users through messaging apps using the large amount of contextual data buried in messages.

Companies are now looking to help users by embedding chatbots inside these message channels to answer questions or assist with various tasks.

Natural Language Interface

The proliferation of devices such as the Amazon Echo has drawn developers toward the idea of a voice-controlled home.

Home appliances are notorious for their clunky user interfaces, and to replace them with smart agents that we could talk to seems like a much better user experience.

Challenges of Building a Successful Chatbot

Define the project scope correctly

- Broad enough for the chatbot to be helpful
- Narrow enough so you're not wasting time building artificial general intelligence.

This means capturing as many user requests as possible, yet still being able to reconcile the nuanced differences between each one.

Example

A travel agency tried to deploy a vacation planning chatbot with a vocabulary base large enough to recognize all the destinations.

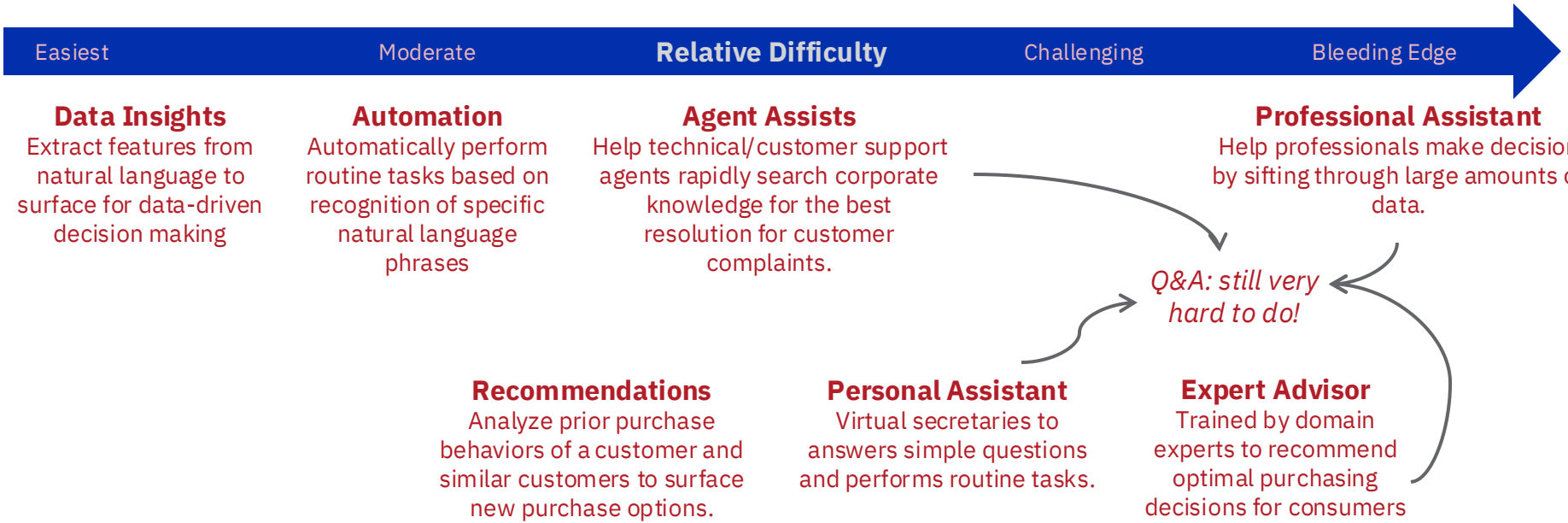
There were over 10 ways people could refer to the Cayman Islands (not counting spelling mistakes).

It took the company months to build a list that could confidently capture all the variations for this one destination.

Best Practices

- Tip #1: Introduce Your Chatbot to First-Time Users
- Tip #2: Add Variations to Your Responses
- Tip #3: Make a Main Menu That's Accessible Anywhere
- Tip #4: Have Context Awareness
- Tip #5: Be Able to Fix Incorrect Inputs
- Tip #6: Handle the "I Do Not Understand" Case
- **Tip #7: Be Careful About Creating a Personality**

Common implementation patterns for conversational cognitive systems



First there were websites, then apps, now bots

One-to-one VS one-to-many

Marketers can no longer rely on the "**yell & sell**" approach and hope for the best. They need to "**chat and listen**."

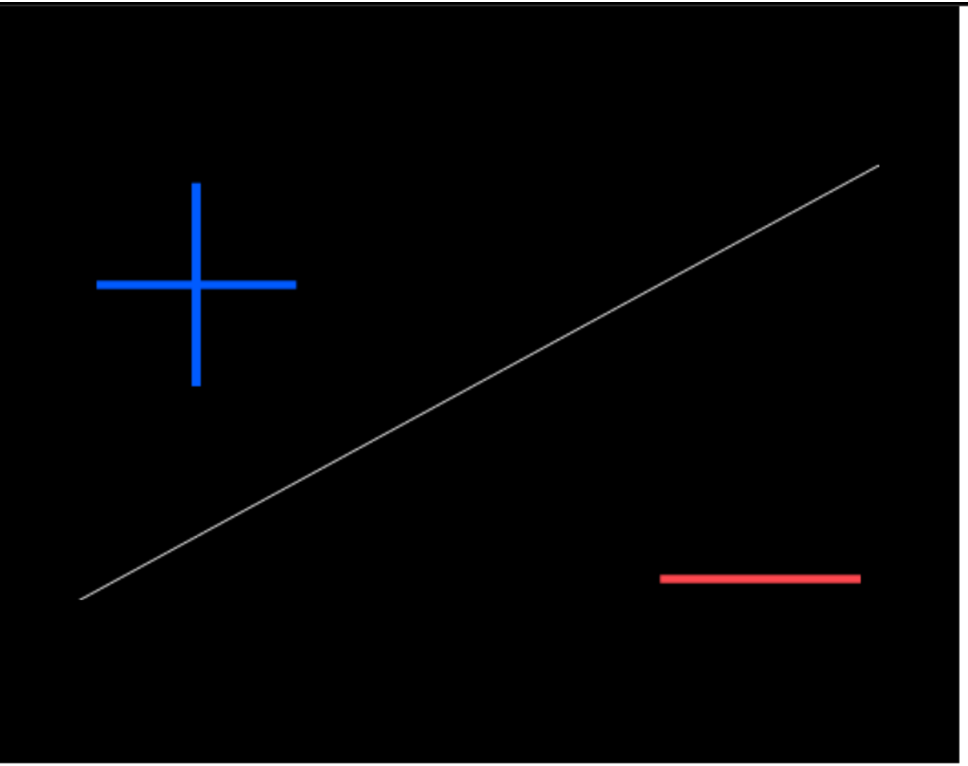
Small data VS big data

One-to-one communication: **snackable data that is actionable**.

Always on VS always perfect

The rise of the "**real and raw**" thanks to live apps like Periscope or Snapchat stands as the new social currency.

THE DEBATER PROJECT



Project Debater

Project Debater is the first AI system that can debate humans on complex topics. The goal is to help people build persuasive arguments and make well-informed decisions.

Watch a live debate

What is Project Debater

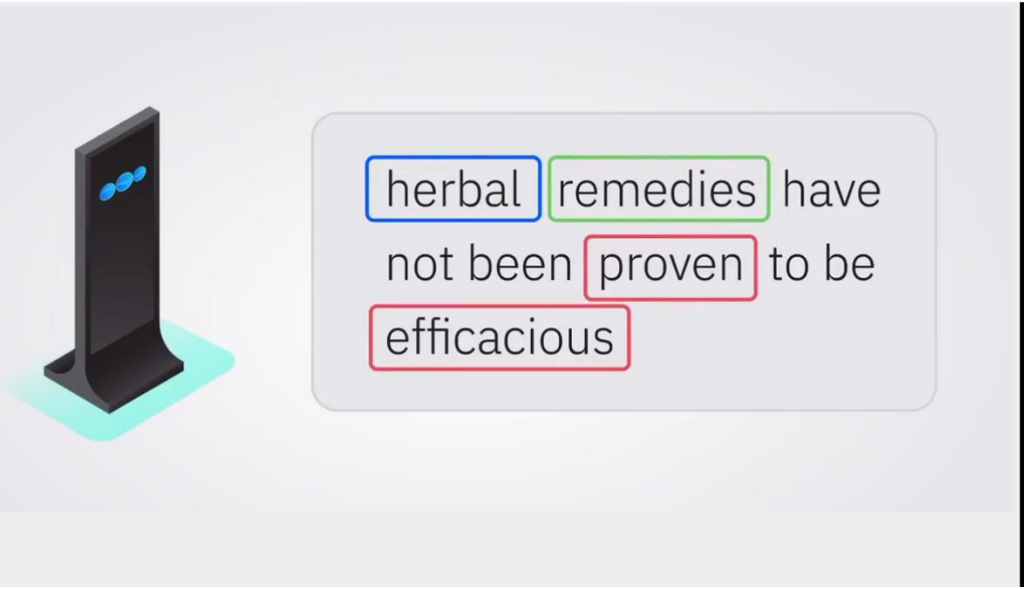
Project Debater is the first AI system that can debate humans on complex topics.

It digests massive texts, constructs a well-structured speech on a given topic, delivers it with clarity and purpose, and rebuts its opponent.

Why teach a machine how to debate

Debate enriches decision making, helping people weigh the pros and cons of new ideas and philosophies.

We debate not only to convince others of our own opinions, but also to understand and learn from each other's views.



Step 1
Understanding a topic

Step 2
Argument construction

Step 3
Content organization

Step 4
Constructing an argument and rebuttal

Step 1
Understanding a Topic

Project Debater's knowledge base consists of around 10 billion sentences, taken from newspapers and journals.

Using AI natural language processing technologies, Project Debater is able to recognize the same concept, even when stated many different ways.

Step 2
Augment Construction

The first step is to build an opening speech to defend or oppose this motion.

Project Debater searches for short pieces of text in the massive corpora that can serve this purpose.

Step 3
Content Organization

In order to debate effectively, the Project Debater needs to construct the strongest and most diverse arguments to support its case.

Project Debater does this by removing redundant argumentative texts, selecting the strongest remaining claims and evidence, and arranging these by theme, creating the base of the narrative to support the motion.

Step 4
Constructing an Argument and Rebuttal

Project Debater pieces all the selected arguments together to create a persuasive speech that lasts approximately four minutes.

The next step is to listen to the opponent's response, digest it and build the rebuttal.