

BADM-4830

# IBM Watson

## Natural Language Understanding

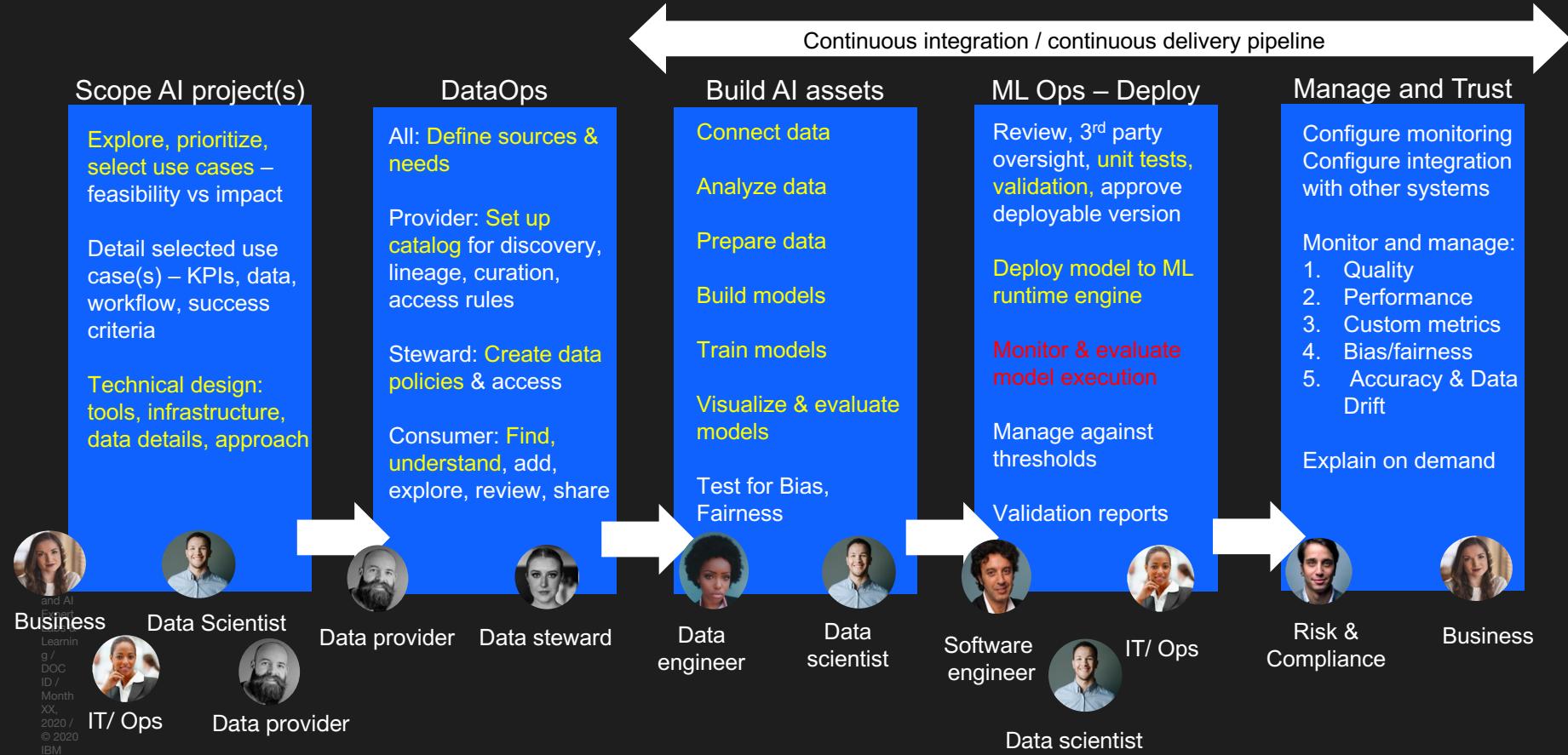


# By the end of this session you will understand ...

- ML Ops workflows
- Module 6 Business case
- Natural Language Understanding
- DevOps best practices
  - Dev environments: Cloud VM vs Local
  - SSH access & keys
  - nodeJS & github refresh
  - Development -> UAT -> Production devOps workflows



# Stages in Operationalizing AI



# How can I apply Natural Language Understanding?

**Business intelligence**: How can I gather business intelligence from unstructured text to create dashboards and reports?

**Social media monitoring**: How do I extract insights from monitoring social media?

**Content recommendation**: How can I recommend content that customers might like?

**Brand management**: How can I know what consumers are saying about my brand?

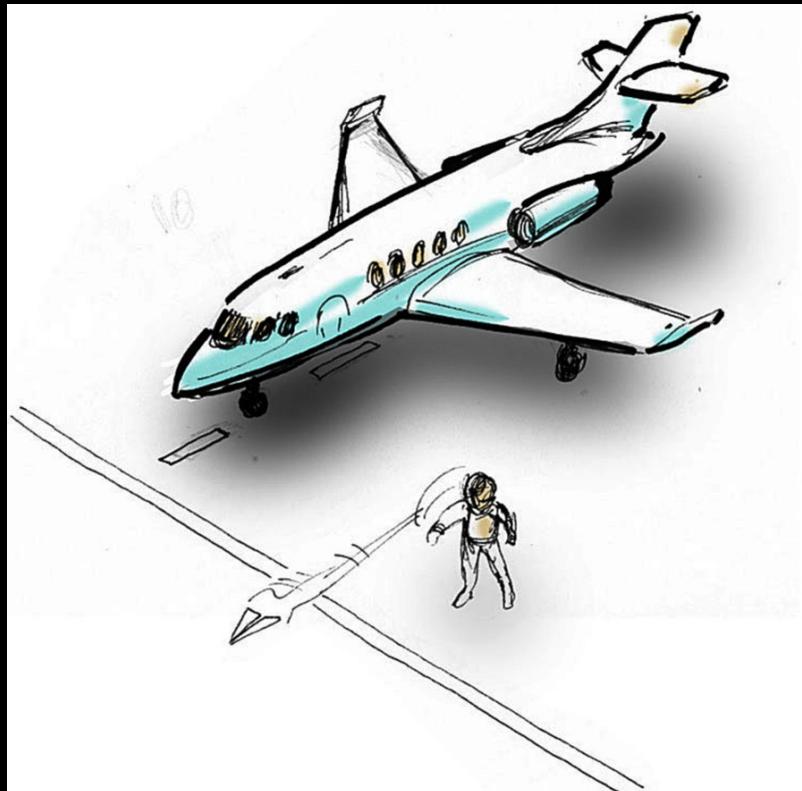
**Advertising optimization**: Where should I place my ads so that I get to the right audience?



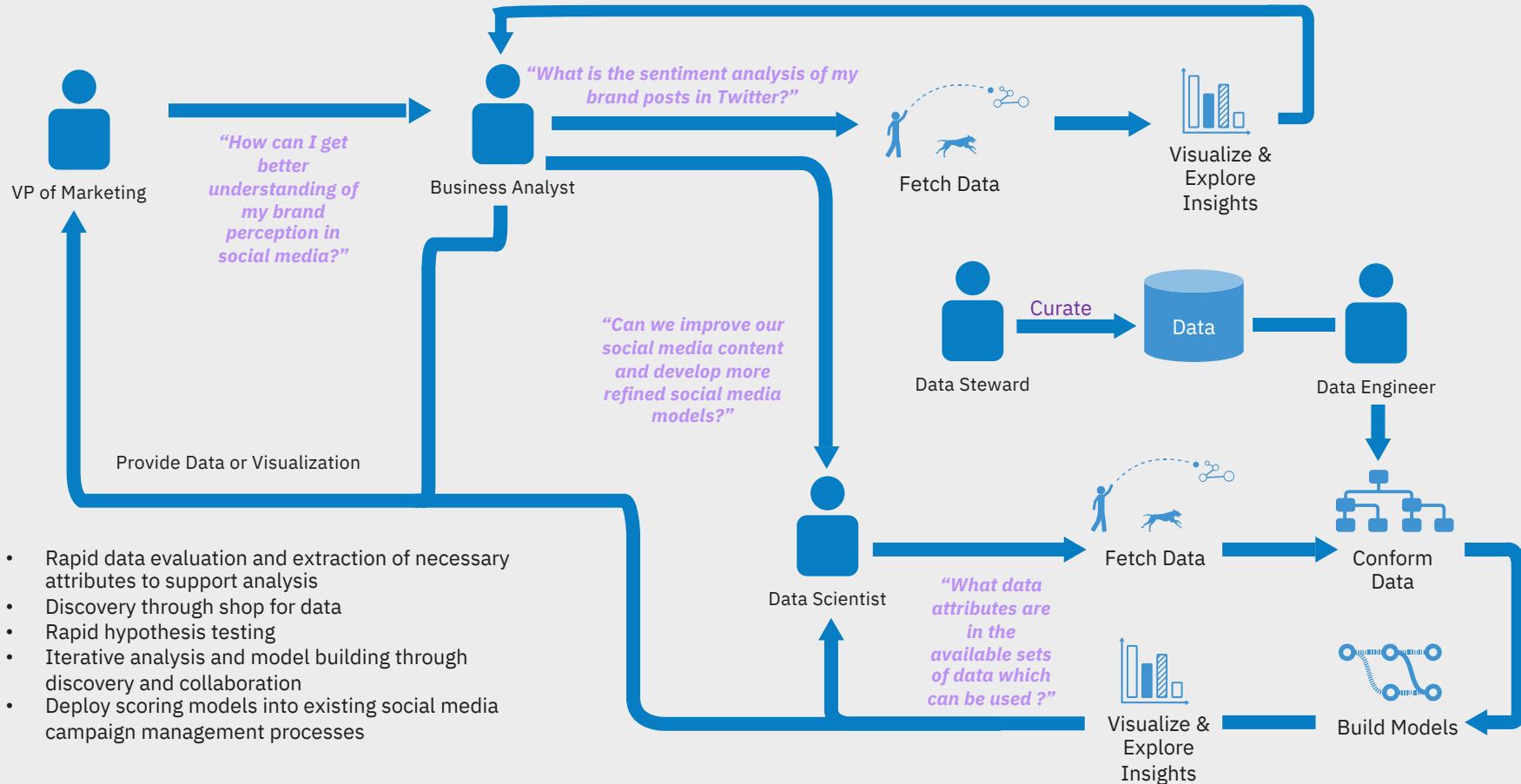
# How can I apply Natural Language Understanding?

**Social media monitoring:** How do I extract insights from monitoring social media?

**Don't overengineer  
a solution if a  
simpler one will  
do.**



# Sample business problem: Social Media Monitoring



# Agenda

## Github

<https://students.yourlearning.ibm.com/activity/URL-64B087D9AA4F>

## Intro to Cryptography

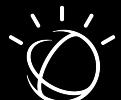
<https://www.ibm.org/activities/intro-to-cryptography>

## SSH keys

<https://docs.github.com/en/authentication/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent>

## Developer Community

<https://w3.ibm.com/developer/docs/gettingstarted/developer-way/>



# Dev VM

Apps + OS      OS Only

 Amazon Linux 2018.03.0.20190826	 Ubuntu 16.04 LTS	 <b>Ubuntu</b> 18.04 LTS	 Debian 8.7
 Debian 9.5	 FreeBSD 12	 openSUSE 15.1	 CentOS 7 1901-01

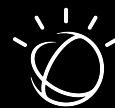
OPTIONAL

Sort by: **Price per month** Memory Processing Storage Transfer

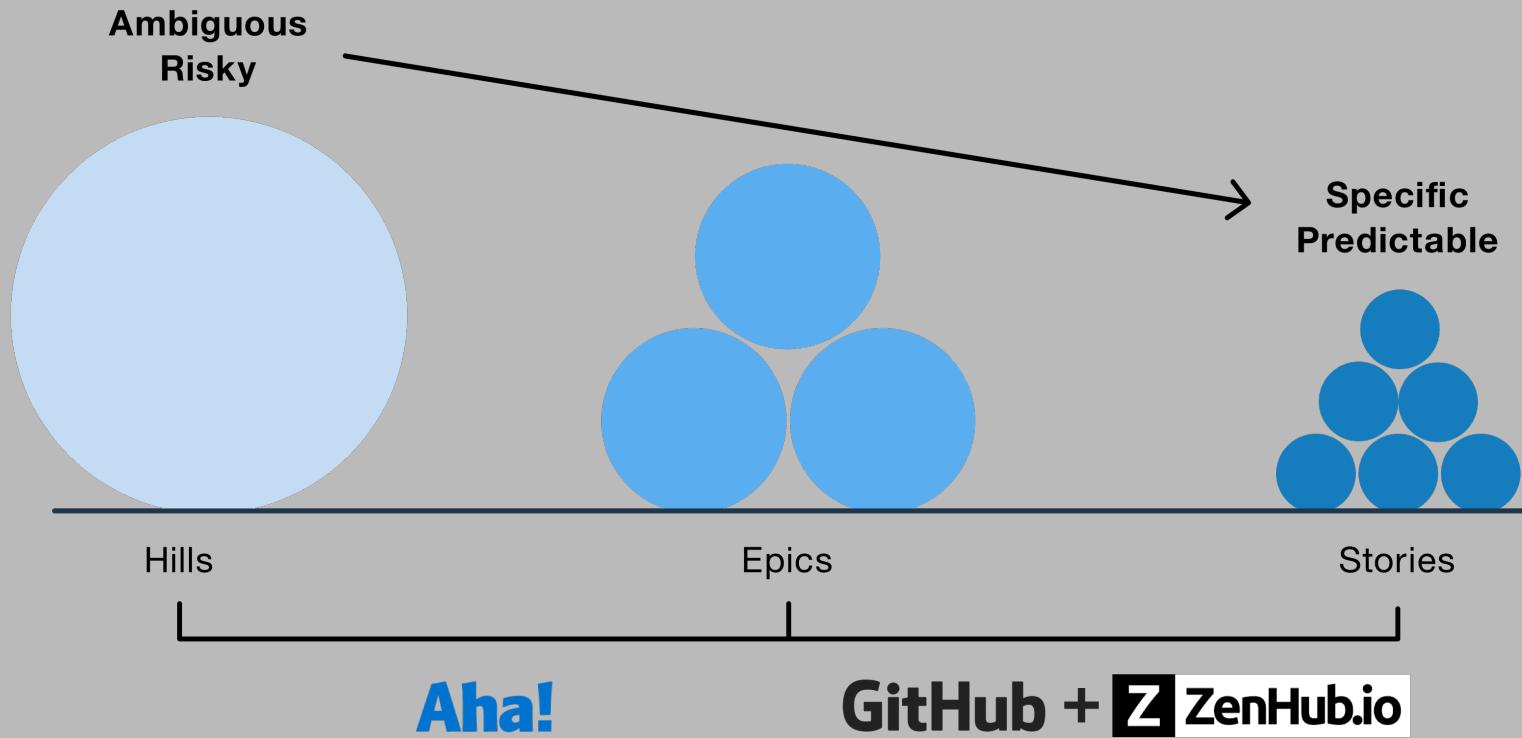
First month free!

 \$3.50 USD	 \$5 USD	 \$10 USD	 \$20 USD	 \$40 USD
--	---	---	--	--

<https://aws.amazon.com/lightsail/>

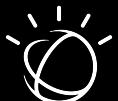


# Ongoing Delivery



# Agenda

- Watson Natural Language Understanding (NLU) overview
- Features of NLU
- Applications & Use Cases of NLU
- Resources



# Watson Natural Language Understanding (NLU)

## Overview



# IBM Watson Natural Language Understanding



- Natural language processing for advanced text analysis
- Analyze semantic features of text input
- A sophisticated suite of natural language processing capabilities to analyze text and extract meta-data from content. This meta data is aka ‘enrichments’.
  - You have the option to customize the enrichments for specific domains via additional training.

keywords

entities

relations

concepts

categories

sentiment

emotion

semantic roles



# IBM NLU

## The solution:

The IBM Watson Natural Language Understanding service enables developers to extract insights from unstructured text to power a new generation of cognitive apps.

Gain rich insights from  
your unstructured data

Saves developers time  
and effort

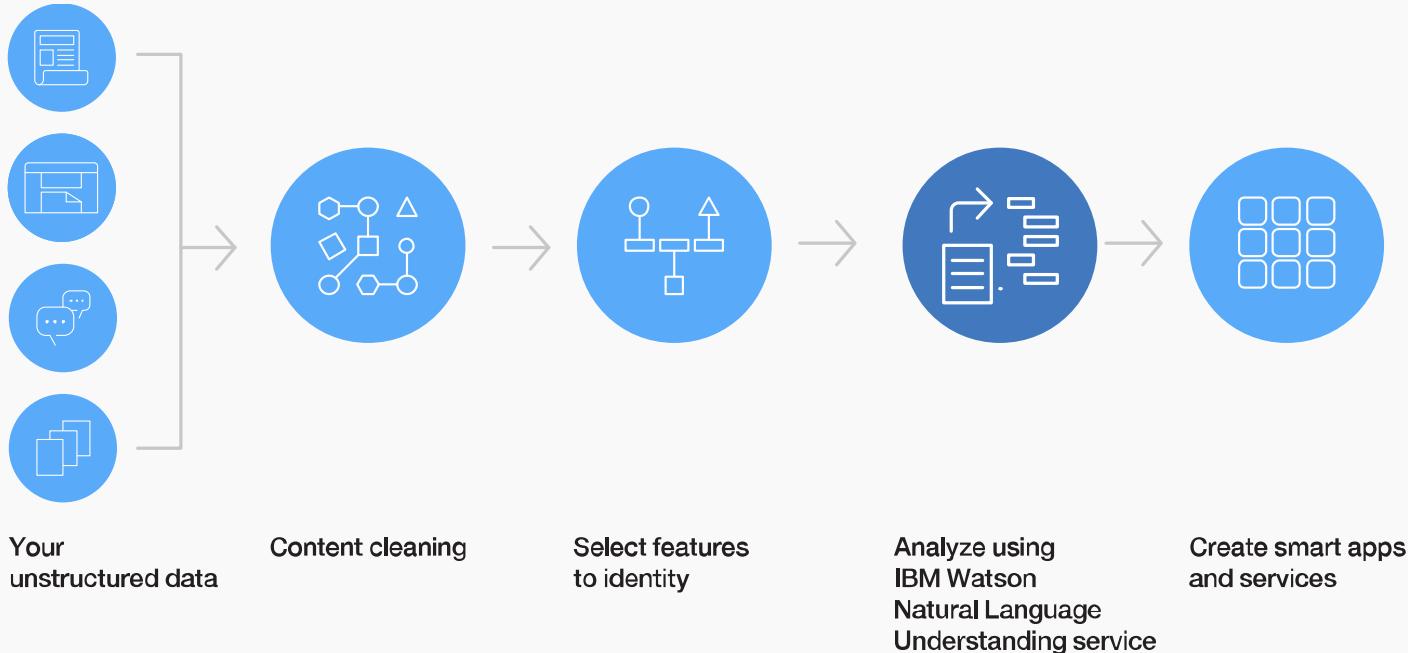
Offers higher accuracy  
metadata extraction

Scales easily for large solutions

No NLP or machine learning  
expertise required



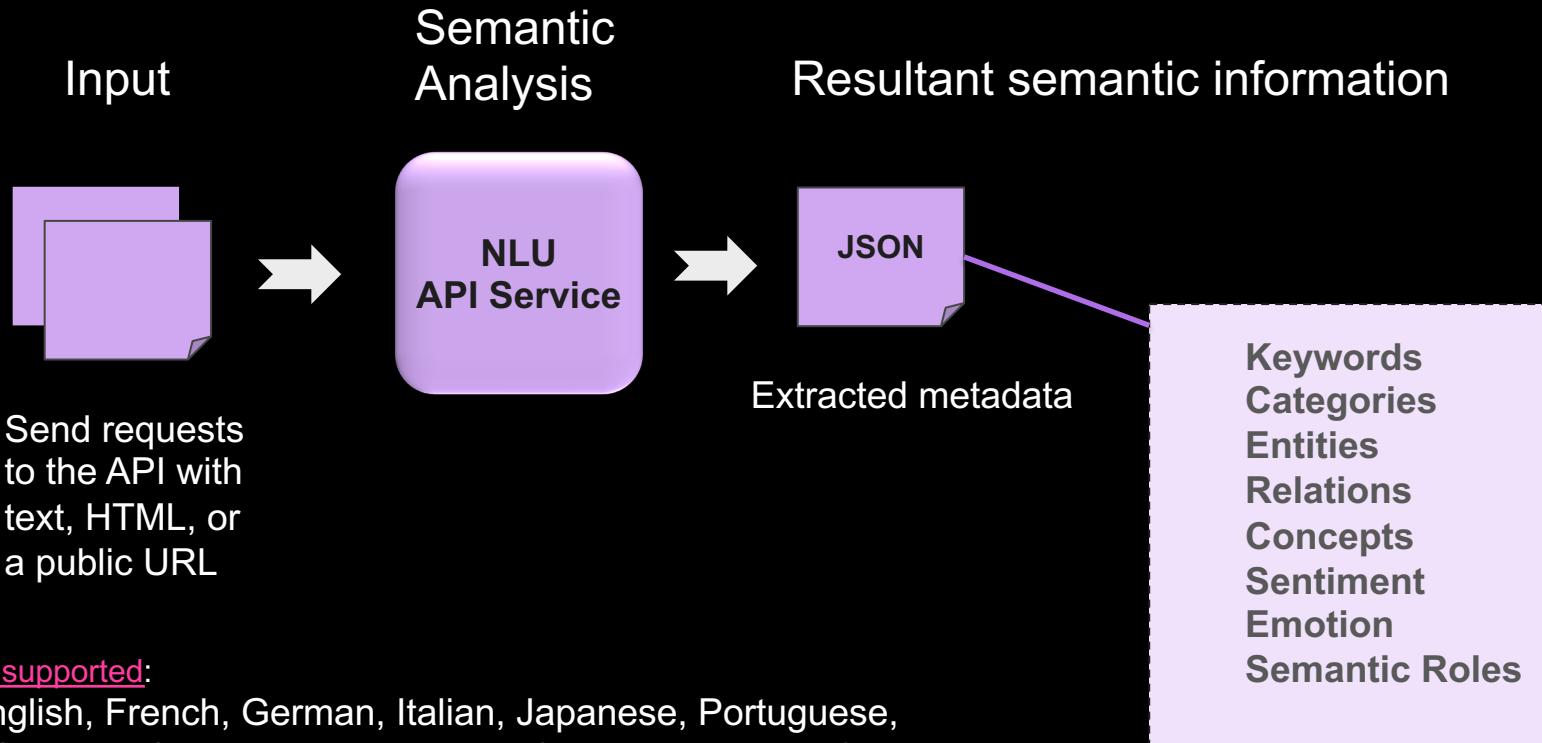
# How it works



IBM Watson Natural Language Understanding makes value-driven decisions easy by giving you the full story behind all of your data.



# NLU Landscape – What it does



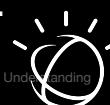
## Languages supported:

Arabic, English, French, German, Italian, Japanese, Portuguese, Russian, Spanish, Swedish, more coming (some variation in feature support across languages)



# NLU Semantic Features to analyze -

- 1. Keywords:** Determine the most important keywords in your content.
- 2. Entities:** Detect important people, places, geopolitical entities and other types of entities.
- 3. Relations:** Identify relationships between entities in your content.
- 4. Concepts:** Identify general (DBpedia) concepts in your content.
- 5. Categories:** Categorize your content into a hierarchical 5-level taxonomy.
- 6. Sentiment:** Determine whether your content conveys positive or negative sentiment.
- 7. Emotion:** Detect emotions such as anger, disgust, fear, joy or sadness that are conveyed by your content.
- 8. Semantic roles:** Identify the subjects of actions, and the objects that they act upon.

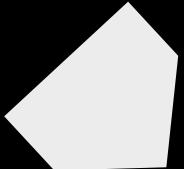


# NLU Semantic Feature - Keywords

## *Input*

A document about the American Civil War:

<http://www.historynet.com/civil-war>



## *Response*

- Civil war
- Union
- Battle
- Confederate

The Civil War  
Facts, Events & Information about The American Civil War: 1861-1865



BATTLES OF BULL RUN

Civil War Facts

Location

Eastern Theater, Western Theater, Trans-Mississippi, Gulf Coast, Sioux Uprising

Dates

1861-1865

Soldiers Engaged

Union: over 2,100,000

Confederate: over 1,000,000

Civil War Casualties

Union: over 250,000

Confederate: over 250,000

See Details of Civil War Casualties

Outcome

Union Victory

Civil War Pictures

The Civil War was the first war that was widely photographed. Many American Civil War Images, pictures, and photos have survived.

See our Civil War Pictures

Civil War Maps

The Civil War made wide use of battle maps.

View our Civil War Maps

Civil War Timeline

See a timeline of events of the Civil War from 1861-1865. See events by year and important Civil War dates.

View our Civil War Timeline

Civil War Battlefields

The battlefields of the Civil War crossed the nation and made famous many previously unknown towns, crossroads, and farms like Antietam Creek, Shiloh and Gettysburg.

View more Civil War Battlefields

More Civil War Facts

VIDEO: Battery H Of The 3rd Pennsylvania Heavy Artillery At Gettysburg



Civil War Times Editor Dave Shoot shows the story of how Battery H of the 3rd Pennsylvania Heavy Artillery found itself in the middle of the Battle of Gettysburg ...



Dan Bullock: The youngest American killed in the Vietnam War

Pic: Dan Bullock died at age 15 in 1969 and efforts to recognize the young African American Marine continue and are highlighted in this Military Times documentary. (Rodney Bryant and Daniel Woolfolk/Military Times)...



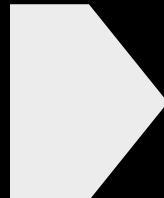
FEATURED

AMERICA'S LOST CITIES

# NLU Semantic Feature - Entities

## *Input*

IBM is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries.

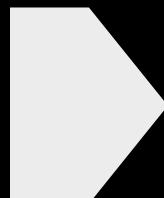


## *Response*

- IBM: Company
- Armonk: City
- New York: State
- United States: Country

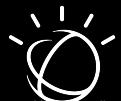
## *Input*

Paris Hilton stayed in a Hilton Hotel in Paris, France



## *Response*

- Paris Hilton: Person
- Hilton Hotel: Facility
- Paris: City
- France: Country

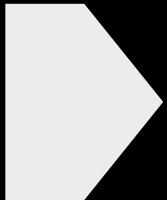


# NLU Semantic Feature - Relations

## *Input*

Text:

Bob Dylan won the Nobel Prize in Literature in 2016. Bob Dylan was born in Duluth, Minnesota.



## *Response*

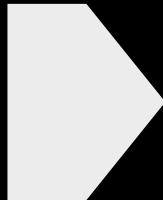
- "affectedBy" relation between "Bob Dylan" and "won"
- "timeOf" relation between "2016" and "won"
- "awardedTo" relation between "Nobel Prize" and "Bob Dylan"
- "bornAt" relation between "Bob Dylan" and "Duluth"
- "locatedAt" relation between "Duluth" and "Minnesota"

# NLU Semantic Feature - Concepts

Identifies high-level concepts that might not be directly referenced in the input text.

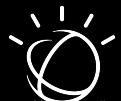
## Input Text:

Machine learning is the science of how computers make sense of data using algorithms and analytic models.



## ***Response*** Concepts tags:

- Computer
- Machine learning
- Artificial intelligence
- Computer science
- Alan Turing
- Scientific method
- Psychology
- Learning



# NLU Semantic Feature - Categories

## *Input*

Text:

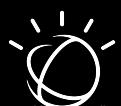
Machine learning is the science of  
how computers make sense of  
data using  
algorithms and analytic models.



## *Response*

These are the responses:

/science/computer science/artificial  
intelligence; score: 0.398614  
/science/; score: 0.386026  
/science/mathematics/geometry; score:  
0.229613

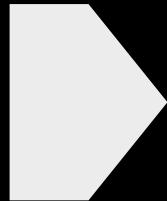


# NLU Semantic Feature - Sentiment

## *Input*

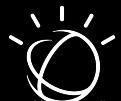
Text:

I'm very upset about the quality of this product.



## *Response*

Negative sentiment (score: -0.890748)

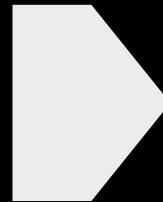


# NLU Semantic Feature - Emotion

Emotion feature detects anger, disgust, fear, joy, and sadness implied in text; analyzes overall emotional tone of the content ; or conveyed by specific phrases.

## *Input*

*I still have a dream. It is a dream  
deeply rooted in the American dream.  
I have a dream that one day this  
nation will rise up and live out the true  
meaning of its creed: We hold these  
truths to be self-evident, that all men  
are created equal.*



## *Response*

Emotion keys & scores (0.0 - 1.0):

- Anger score: 0.06
- Disgust score: 0.07
- Fear score: 0.04
- Joy score: 0.51
- Sadness score: 0.35.

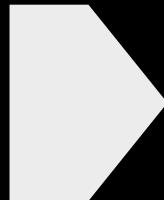


# NLU Semantic Feature – Semantic Roles

Parse sentences into subject-action-object form, and identify entities and keywords that are subjects or objects of an action. For example:

## *Input*

Text: "In 2011, Watson  
competed on Jeopardy!"



## *Response*

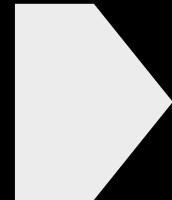
These are the responses:  
Subject: Watson  
Action: competed  
Object: on Jeopardy



# NLU Semantic Feature – Language detection

## *Input*

Input is from a web page or text, for example:  
<https://www.ibm.com/ares/>



## *Response*

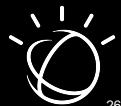
Here is the response:

```
{  
  "retrieved_url": "https://www.ibm.com/ares/",  
  "concepts": [  
    {  
      "text": "IBM",  
      "relevance": 0.965628,  
      "dbpedia_resource":  
        http://es.dbpedia.org/resource/IBM  
    }  
  ],  
  "language": "es"
```



# NLU Highlights

- 1. Comprehensive NLP stack:** A rich set of enrichments to extract metadata from unstructured text for deep analysis in various use cases
- 2. Broad Language Coverage:** Depending on the feature, NLU can understand text in nine+ different languages
- 3. Scalability:** Cloud-based service with the infrastructure to support several NLU enrichments per month
- 4. Custom models:** Use IBM Watson Knowledge Studio to collaborate on the creation of custom annotation models



# How you can use Watson Natural Language Understanding

## Applications and use cases



# How can I apply Watson Natural Language Understanding?

**Business intelligence:** How can I gather business intelligence from unstructured text to create dashboards and reports?

**Social media monitoring:** How do I extract insights from monitoring social media?

**Content recommendation:** How can I recommend content that customers might like?

**Brand management:** How can I know what consumers are saying about my brand?

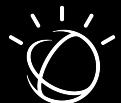
**Advertising optimization:** Where should I place my ads so that I get to the right audience?



# Hands On Lab

## Watson Natural Language Processing

w/ IBM guidance



# Let us check this out !!

<https://natural-language-understanding-demo.ng.bluemix.net/>



# Resources

## Documentation

- <https://console.bluemix.net/docs/services/natural-language-understanding/getting-started.html#getting-started-tutorial>

## API Reference

- <https://cloud.ibm.com/apidocs/natural-language-understanding/natural-language-understanding>

## Download the SDKs

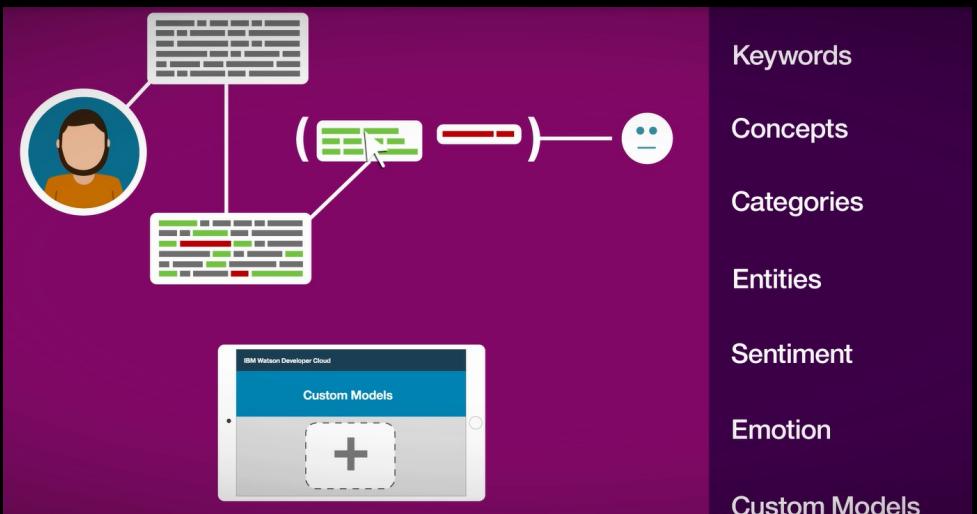
- <https://github.com/watson-developer-cloud>

## Demo Application

- [https://ibm.biz/nlu\\_demo](https://ibm.biz/nlu_demo)

## Videos

- [https://ibm.biz/nlu\\_videos](https://ibm.biz/nlu_videos)



The screenshot shows the Natural Language Understanding Service demo application. It features a text input field containing a sample sentence about a sunken treasure ship. Below the text input is an 'Analyze' button. To the right, there is a large graphic for the 'Natural Language Understanding Service' featuring a smiling woman and the Watson logo.



# Methodology

## Watson Natural Language Understanding



# Overview

Analyze/  
Review/  
Adjust the  
Model

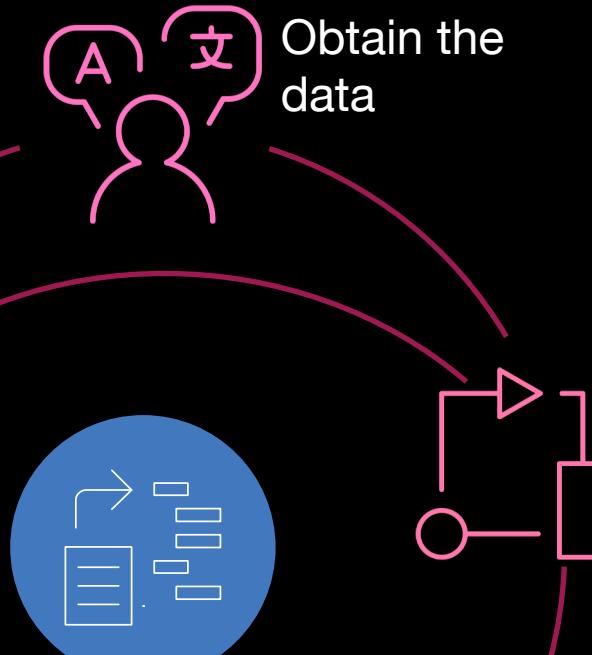
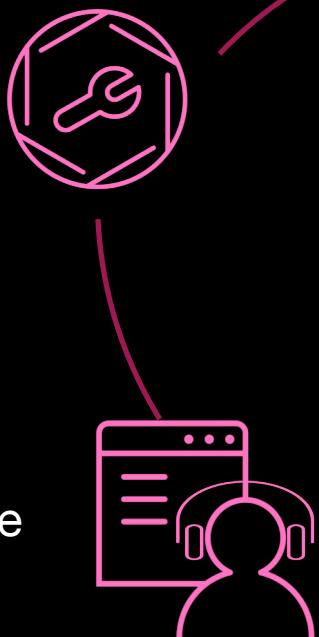
Customize the  
Model

Natural  
Language  
Understanding

Obtain the  
data

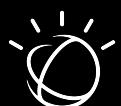
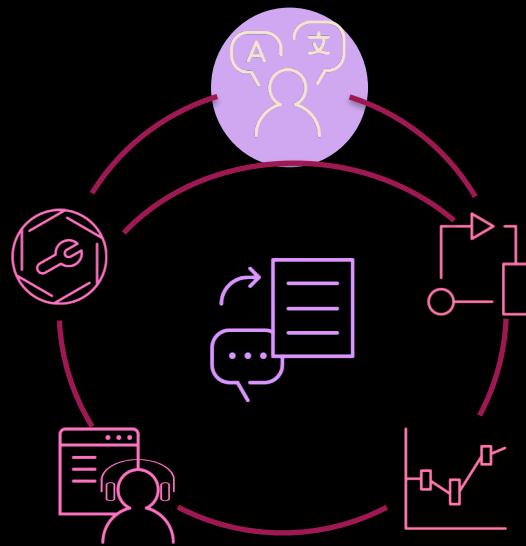
Content Cleanup

Select  
Features



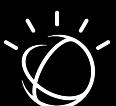
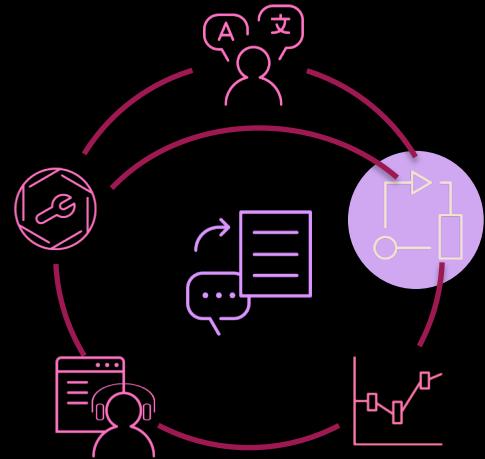
# 1- Obtain the data

- What is my solution ?
- Who is my end user?
- Find out the source of input data?



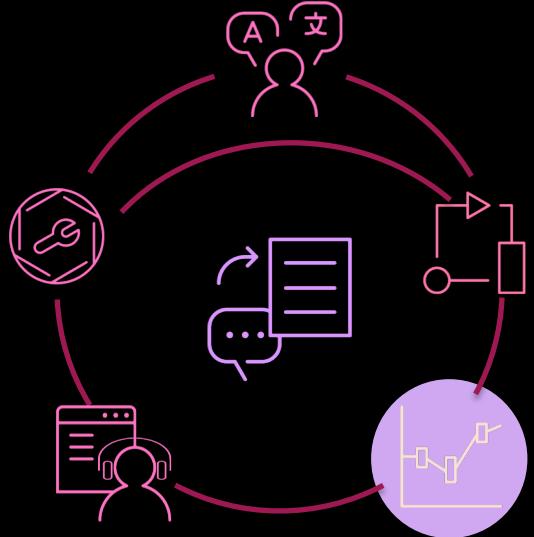
## 2- Content Cleanup

1. Content cleanup is totally dependent on the use case. This is an optional step
2. Unstructured data is coming from various sources so sometime content cleaning is necessary
3. Make sure not to eliminate important information



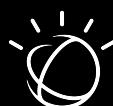
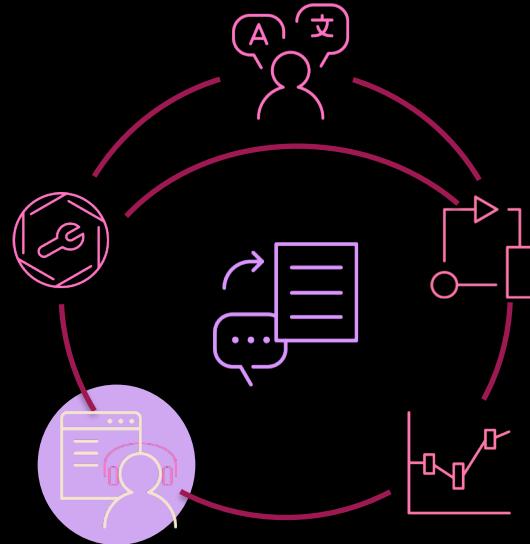
# 3- Select Features

1. Select NLU features as per the use case
2. Don't do analysis for unnecessary features
3. Decide if you need to customize the model



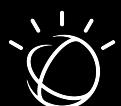
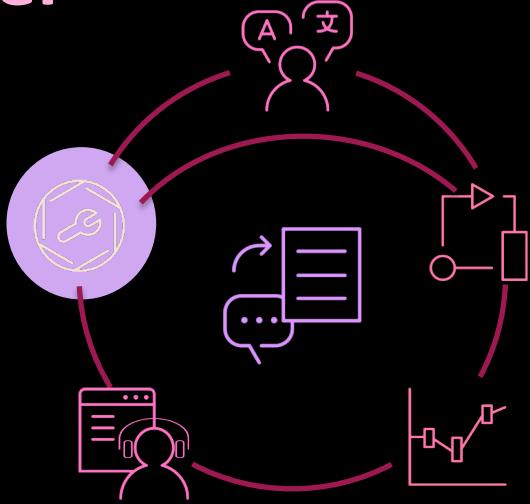
# 4-Customize the Model

1. Model customization is necessary to analyze the content as per organization terminology
2. Engage subject matter experts for model customization



# 5- Analyze/Review/Adjust the Model

1. Analyze the model to generate the customer insights
2. Review the results
3. Adjust the model if necessary



# 6- Repeat

- Repeat the process to improve the result

