Unit 2

Cognitive Computing

Syllabus

- Foundation of Cognitive Computing
- Design Principles for Cognitive Systems
- Natural Language Processing in Support of a Cognitive System
- Representing Knowledge in Taxonomies and Ontologies
- Applying Advanced Analytics to Cognitive Computing
- The Process of Building a Cognitive Application.

Self-learning Topics: Cognitive Systems such as IBM’s Watson.

Cognitive computing

- To use of **computerized model** to simulate the human thought process in complex situation in case the answer may be ambiguous or Uncertain
- Analog to human brain to analyse in context all types of data in DB or KB to Unstructured data in text, image, voice, sensor, and video
- 3 fundamental principle
  - Learn
  - Model
  - Generate Hypothesis

Cognitive computing

A practical approach to computing applications using the science of minds

- **Cognitive computing:** an approach to support human problem by solving with learning machines that will detect new discoveries .
  - **Represents a new way of creating applications** to support business and research goals.
  - **Many technologies that have matured** enough to become commercially viable
  - **A data driven technique** to manage, understand and **analyse unstructured data** in context to the question asked to the DB/ Knowledge Base.

  *In most of the organization 80% of data that is collected and stored as unstructured data.

Overview of Cognitive Computing

3 important concept help make a Cognitive System

1. Contextual i**nsight** form model
2. **Hypothesis** Generation
3. Continuous Learning of data across time

Overview of Cognitive Computing

3 fundamental principles :

- *Learns:* A cognitive systems learns based in training and observation .
- *Model:* to learn the system needs to create a model or a representation domain and assumptions dictate what learning algorithm are used to score the hypothesis or discover new insights.
- *Generate Hypothesis* : cognitive system belies there is no single correct answer but its probabilistic. A CS uses continuously changing inflow of data to train, test or score a hypothesis .

Applications

- Systems enabling , Rerouting Traffic, City Manager Applications, analysing disrupted Weather events .
- **Medical cognitive healthcare development rich in text based data source (**symptoms, disorders or disease and EMR should be considered from various practitioners) successful patient outcome, to aid the doctors to a better understanding of treatment options through continuous learning.
- **Smart city applications** to control pollutions, improve traffic congestions and fight online crimes**.**

Driving Factors for Cognitive Systems

1. The growth of **complex volumes** of unstructured multimedia Corpora due to machine learning , intelligent smart assistants, multimedia files and precise Sensing devices.
2. **Increasing innovative technology** for a flexible pricing rates for computation and storage
3. In-depth emerging research across globe on collaborative learning and competitive rapid discoveries.

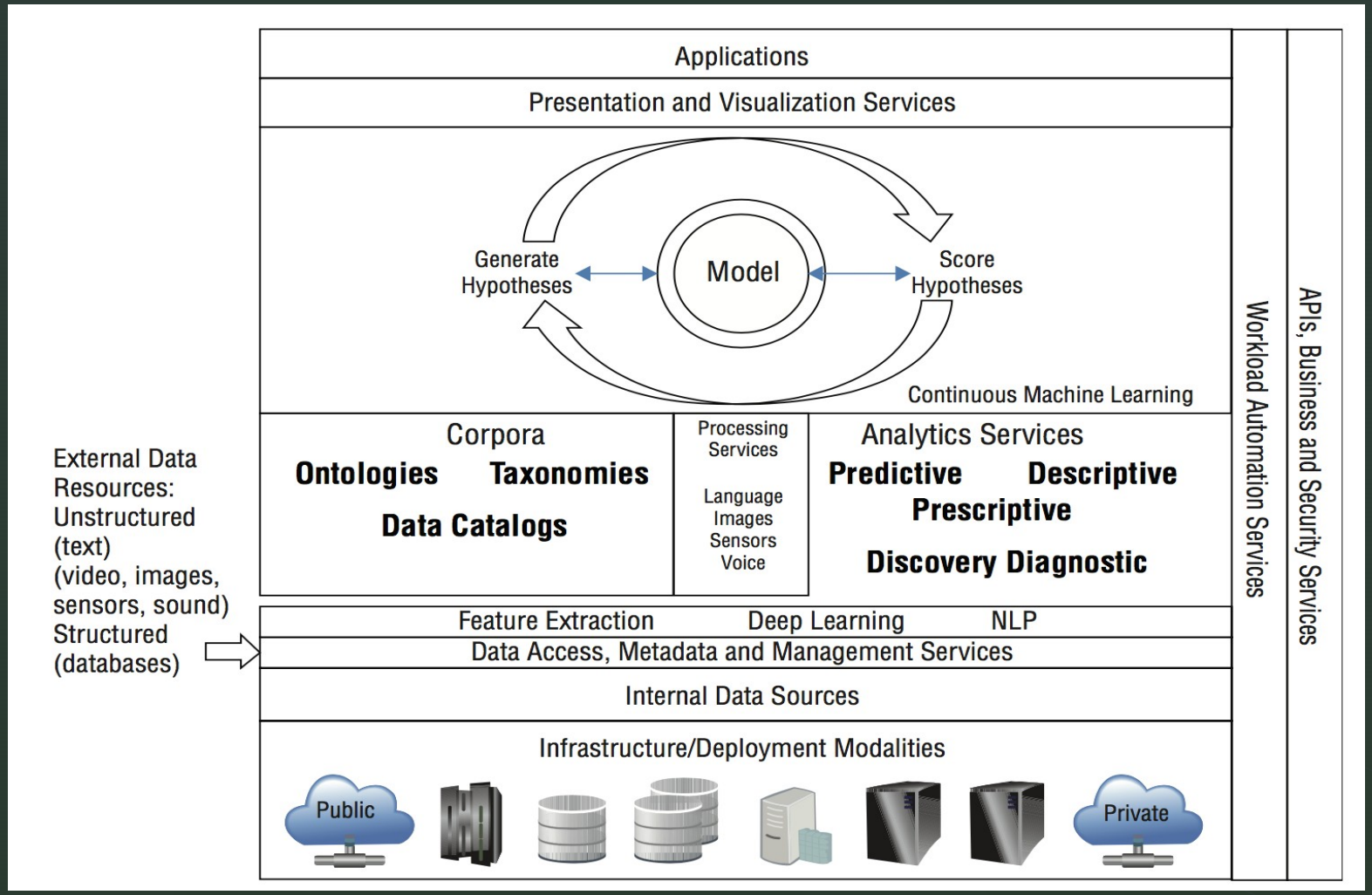
.

Foundation of Cognitive Computing

Features of CS :

- **Learn from experience** with data / evidence and improve its own knowledge and performance without reprogramming .
- **Generate or evaluate** conflicting hypothesis based on the current state of its knowledge .
- **Reports on findings** to justify its conclusion based on confidence and evidence .
- **Discover pattern** in data with or without guidance, Use a variety of predictive analytic algorithm and statical techniques .
- Emulate process or structure (memory management, knowledge organization process, or modelling the neuron) found in NL systems
- **Use NLP** to extract meaning from textual data and use dl tools to extract features from image, video, voice and sensors.

Elements of CS [1]



List of Designs Principles for Cognitive Systems

Cognitive system are designed to learn from their experiences with continuous inflow of data at a time frequency .

1. Data, Characteristic defining Data & Optimized data
2. Generate Hypothesis
3. Ability to extract features using ML algorithms /models
4. Perform experimentation /Analytics
5. Validate Hypothesis
6. Continuous learning without reprograming
7. Hypothesis Scoring
8. Presentation & visualization service

Design Detail in Detail

- The design needs to support different characteristics of data Access, Manage and Analyse data in context.
- Generate and score hypothesis on basis of accumulated knowledge.
- The system continuously updates based on user interaction and data and becomes smarter over time in an automated way .

Data and Data MGMT

a. Data considered in CS is always dynamic
b. Identify and apply some policy for protection, security and compliance to state ,federal and international legislation .
c. The external and internal sources of data have to be considered to solve a specific domain problem .
d. How to optimize the organized data (aggregation) to increase the efficiency of search and analysis .
e. How to integrate data across multiple corpus (ontology) .
f. Unreliable evidence should be deleted otherwise it might divert the outcome
g. Updating of data source, at what frequency ?

Feature Identification, Abstraction & Extraction

1. Any data that is unstructured from video ,image to text must be processed in this layered to find the underlying structure .
2. Ability to identify the correct features and then abstraction of data is done to support the Machine Learning (ML).
3. **Feature extraction** and DL refers to a collection of technique, used to transform data into representation that captures essential properties to more abstract form that can be processed in ML algorithm .

Performing analytics

Analytics provides an insight and guides towards some action or decision .

a. A wide range of standard analytical component available for descriptive predictive and perspective task within statistical packages or component libraries .

b. A number of package algorithms such as regression analysis are used within solutions .

Continuous learning without reprograming- ML

1. **Continuous learning without reprograming** i.e. ML is the heart of cognitive system , to identify pattern in data a typical machine learning algorithm looks for pattern in data and takes or recommends some action based on what it finds.
2. ML uses inferential statistics(predictive) techniques, finding the right algorithm to discover the unknown and known patterns , association between data elements to solve the problem . how to exploit these patterns manually i.e. supervised or unsupervised learning

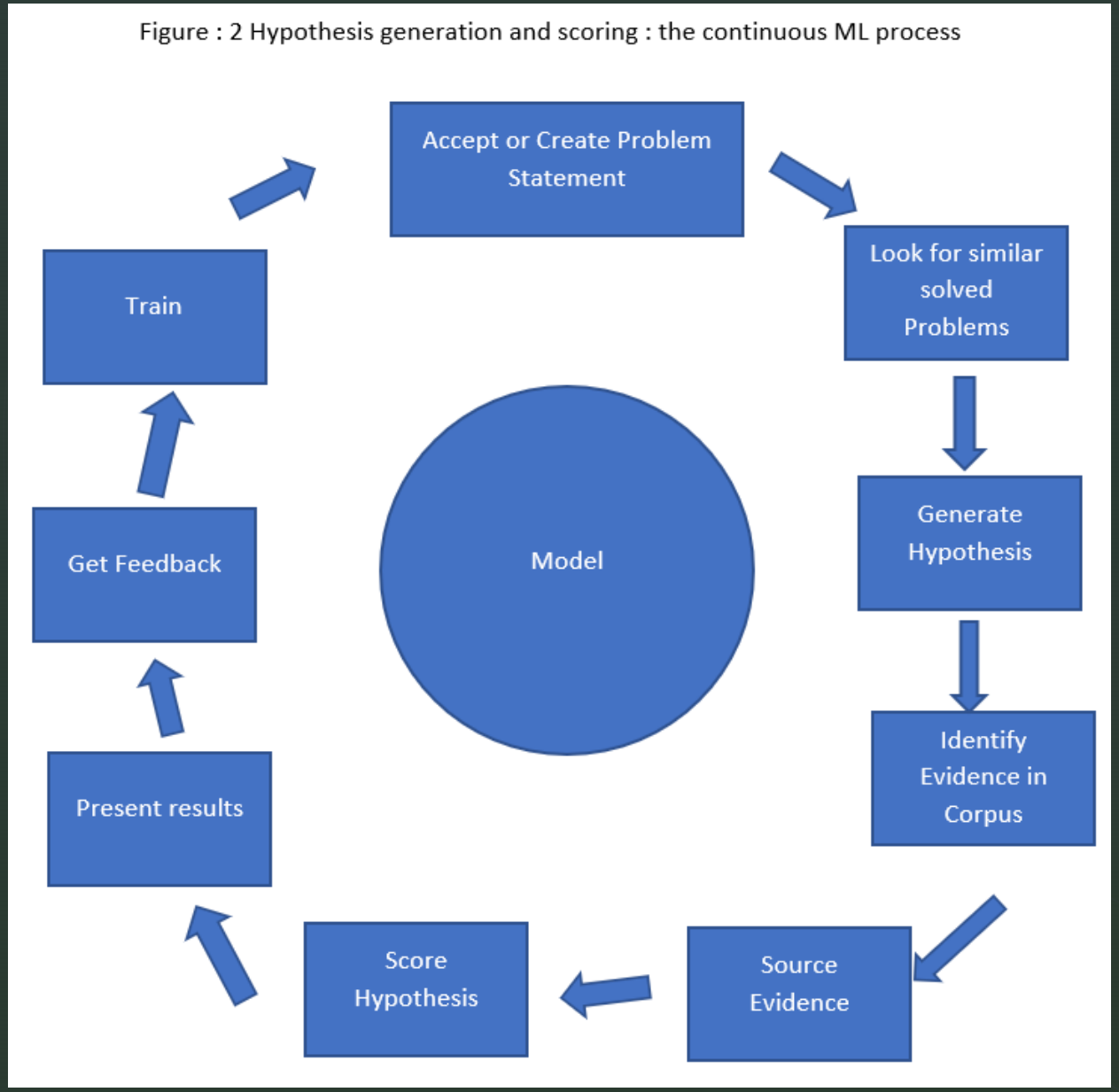
Hypothesis generation and scoring

Hypothesis is a testable assertion based on the evidence that explains some observed phenomenon or relationship between elements of with.in domain. A key concept that support evidence or knowledge for a causal relationship . “if P then Q” , where P is the Hypothesis and Q is the outcome .

a. We conduct experiment to test the **hypothesis is formulated** to answer a question about a phenomenon based on evidence .the goal to use hypothesis sis to detect the cause effect relationship .

b. **Hypothesis Scoring** : To evaluate the date to test whether the evidence exists to support the hypothesis .the actual weights are associates with each piece of supporting evidence are adjusted based on the experience or feedback of the system

Figure : 2 Hypothesis generation and scoring : the continuous ML process



Presentation & visualization Infrastructure service

- Reports should convey the findings, so that user may convey the data in an unambiguous way .
- Infrastructures required are Distributed and parallel system .

NLP in support of a Cognitive systems

- *Examples* :
  - In retail market there are billions of conversations that are leading indicator of future trends .
  - The information buried inside the video; audio files can impact various fields.
- **Tools** for this process of I. identifying the meaning of the individual words include categorization, thesauri, ontology, tagging, catalogues, dictionaries and language models ii. to ask a question to get a meaningful answer from Knowledge base, structured database, query engine.
- **NLP technique** interpret and enables the CS , for a Repository of documents , Reports, e-mails messages, Speech recognition, Images, Videos can be analysed to make good decision .
- These **unstructured information** as in text form needed to be converted to paragraph, sentence and words, to extract the meaning, the relationship between the massive amount of NL Corpus / elements /words and to understand the sentiments.
- **To get the answer** to various questions posted by the user to CS such as
  - Is there a data ? when was the text generated ? (Are they reference to any time or place)?
  - Who is speaking? Are there pronouns in the text? to whom or what they refer to ?
  - Is there important information in previous paragraph ? Are Is there reference to other document in the text ?
  - And to solve ambiguity using the dictionary DB .

NLP in support of a Cognitive systems

- *Understanding linguistics:*
  - To communicate with computer through stages of NLP
  - Language identification and tokenization and phonology(sound), Monology(structure of words)
- *lexical analysis :*
  - Connect each word with dictionary meaning
  - Tagging n-gram words to determine the frequent subset of words
- *Syntactic analysis :* parse the sentence based on given grammar rules.
- *Construction of grammar :* meaning of the content , the intension and extends as pragmatic
- *Disclosure analysis :*coherence of connected sentence and meaning
- *Hidden Markov Models (HMM) :* Statistical model for processing both image and speech understanding are Markov model .

Examples **Applying NL technologies to Business problems**

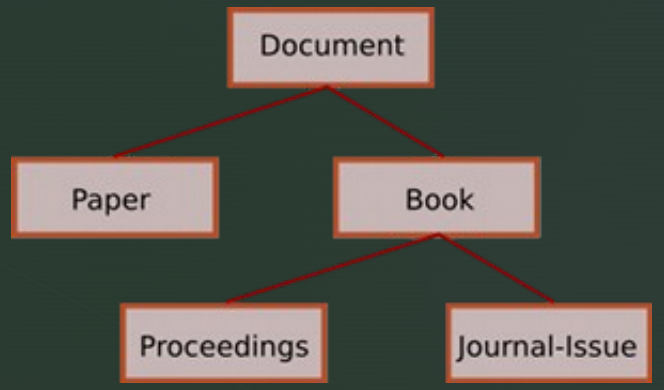
- Enhancing Online Shopping Experience: through search keywords , online chatbot
- Leveraging the connect world of IoT: Re-Routing Traffic under circumstances
- Voice of Customer: understanding the voice of customer and their saying , sentiment analysis
- Fraud detection : detecting internal and external threats DB, from stealing intellectual properties

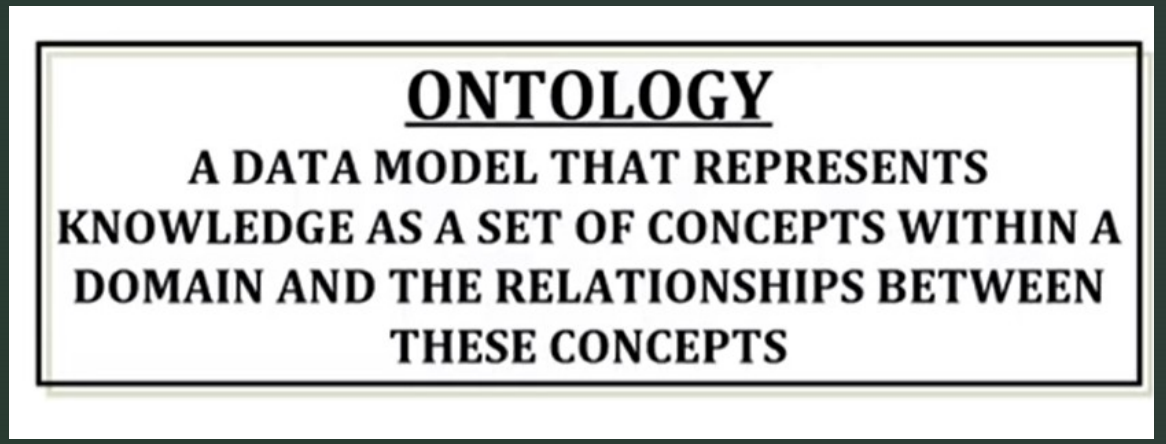
Standard or Models for knowledge representation in an organizational structures with relation , rules or properties or attributes and categorization

Representing Knowledge in Taxonomies and Ontologies

Taxonomy

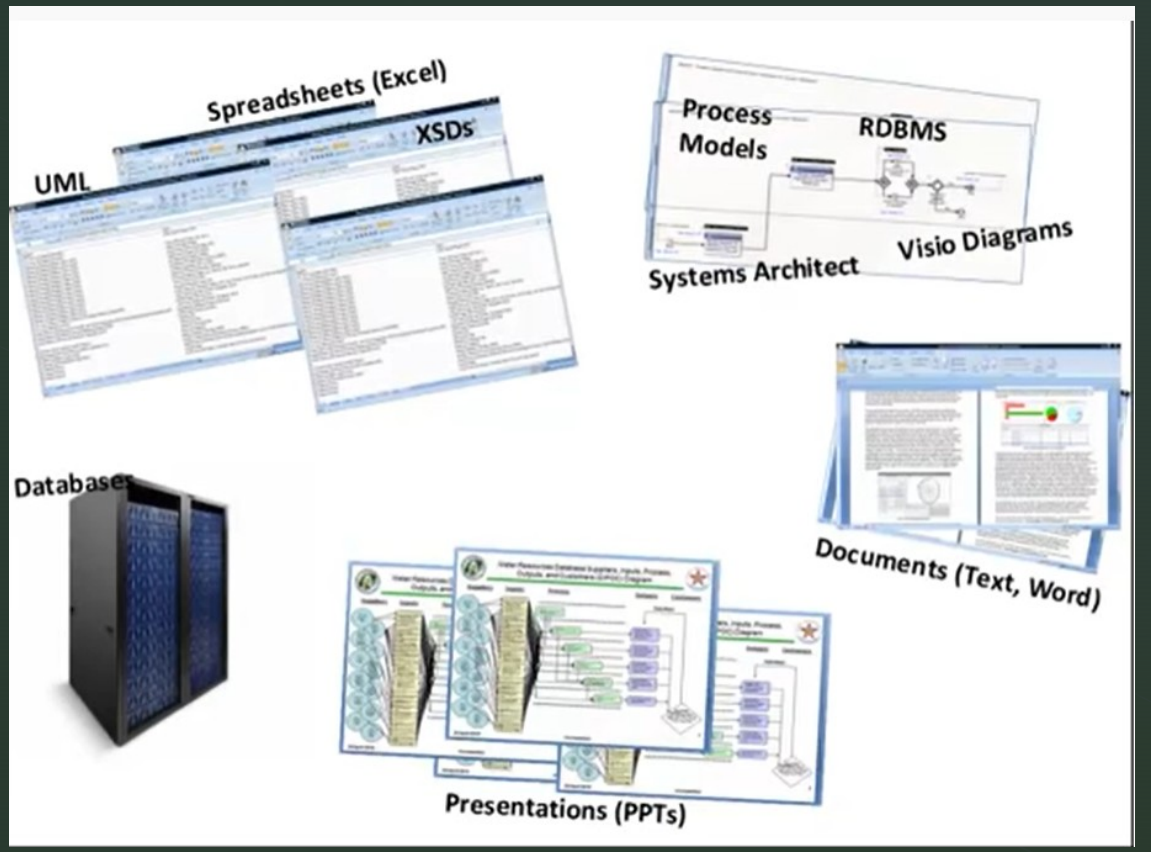
Basic classification system in an hierarchical fashion the provides names to object and their relations



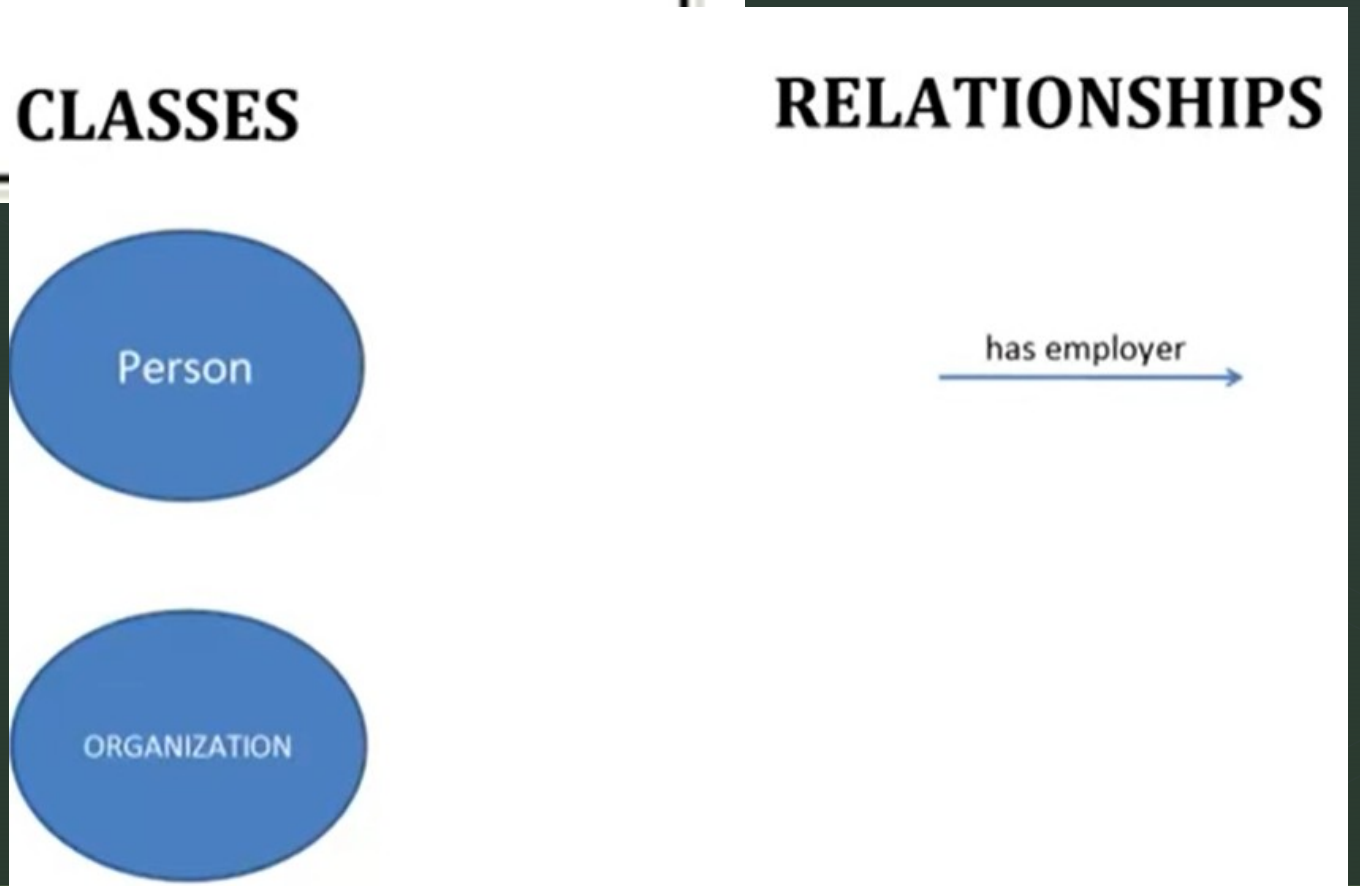


Ontology

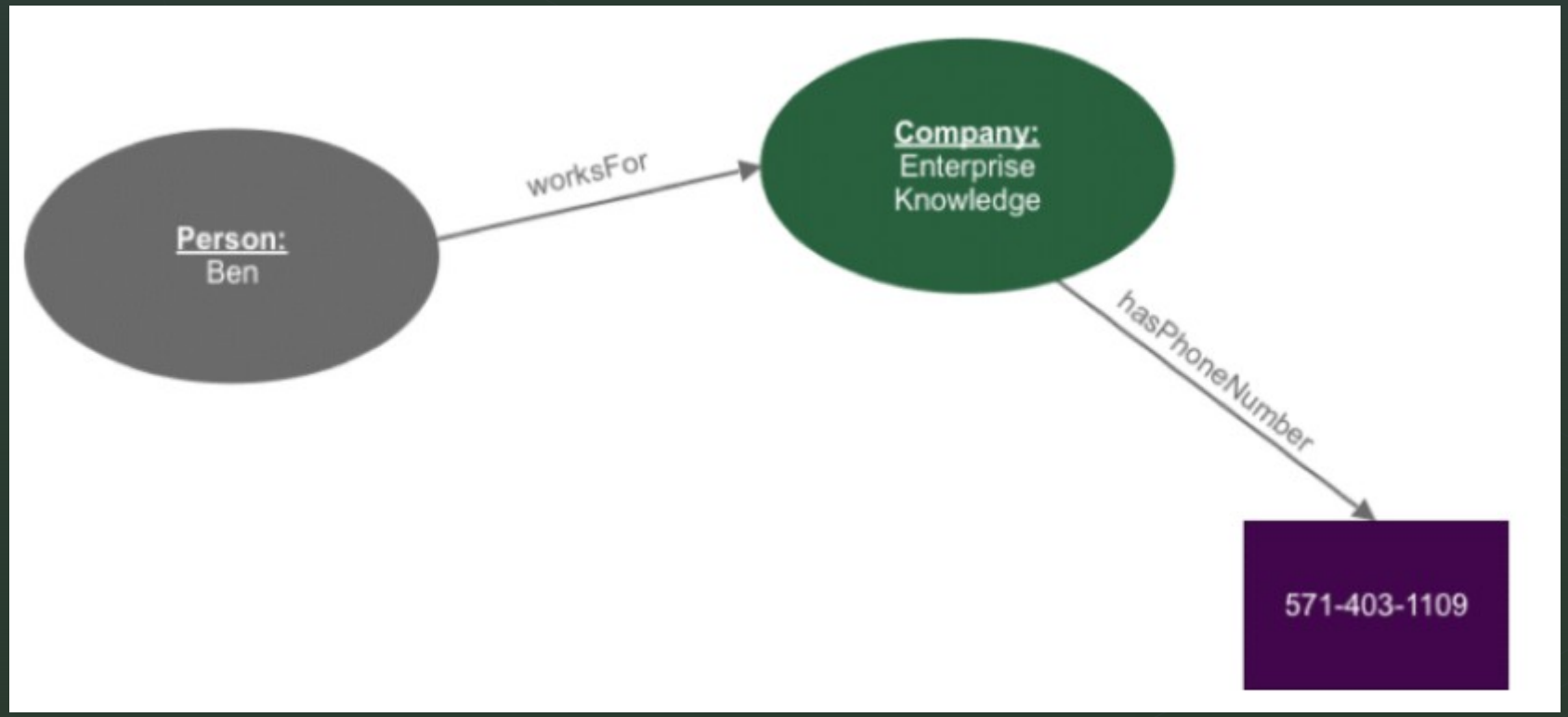
Some common use cases include ontology modelling for content categorization, text mining, user experience, and artificial intelligence


Spreadsheets (Excel)
XSDs
UML
Process Models
RDBMS
Visio Diagrams
Systems Architect
Database
Documents (Text, Word)
Presentations (PPTs)

ONTOLOGIES
ARE MADE UP OF CLASSES AND RELATIONSHIPS



Ontology Component	Definition	Example
Class	Type of entity	People, Places, Companies, etc.
Relationship	Link between objects in the ontology	Ben worksFor Enterprise Knowledge
Attribute	Data associated with an object in the ontology	Enterprise Knowledge hasPhoneNumber 571-403-1109

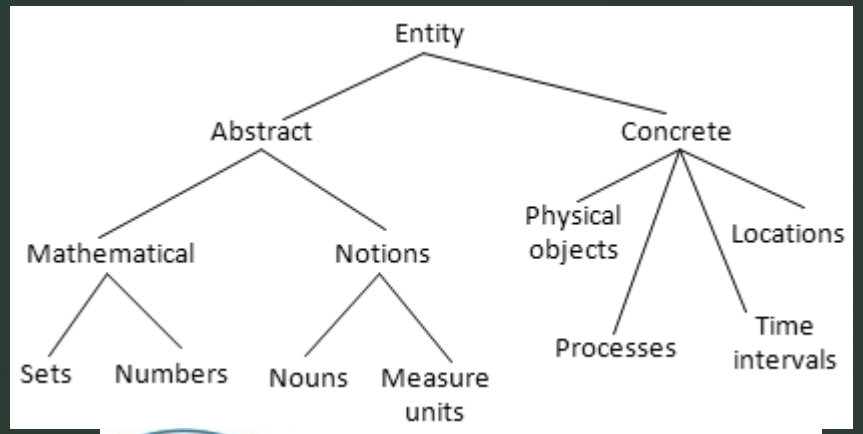


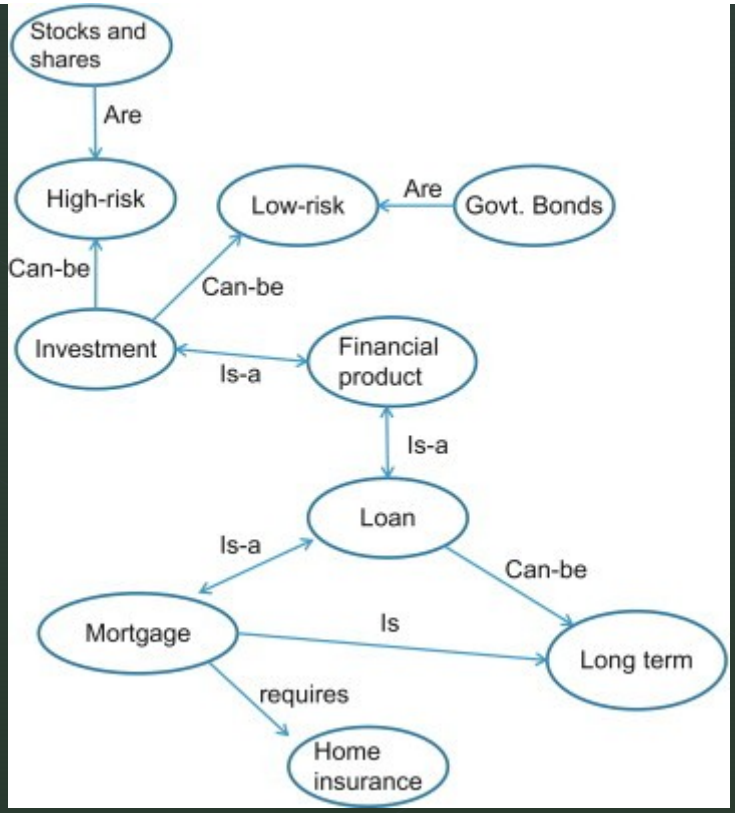
1. **select the use case : a domain**
2. **Extract Classes:** Ontology classes are ***high level categories that encompass the concepts in a given domain***.
3. **Adapt Existing Taxonomy :**Using the classes extracted from Step 2 we can begin to adapt the original taxonomy for ontology transformation. This process involves. moving concepts to fit hierarchically beneath each class
4. **Incorporate Semantics:** the taxonomy has been updated and ready for ontology transformation, it is time to incorporate semantics.
5. **Create Class Hierarchy :** A class hierarchy includes a class and sub classes organized hierarchically a Subclass ***inherits*** the ***attributes*** and ***relationships*** of the parent class.
6. **Rinse and Repeat :**
   - Ontology Engineering is an ***iterative process***. Ontologies should regularly be ***validated*** against initial use cases. Expect the ontology to ***evolve*** and ***extend*** overtime

https://enterprise-knowledge.com/from-taxonomy-to-ontology/

Other methods of knowledge representation

1. Simple tree
2. semantic web





Applying Advanced Analytics to Cognitive Computing

- Collection of techniques and algorithm for identification of pattern in large complex high velocity data set with varying degrees of structures .
- includes statistical models ,predictive analysis, machine learning, neural network, text analysis and other advance data mining techniques
- statistical tech is used in decision tree, linear and logistic regression social network analysis time series analysis .
- Analytic process helps to detect anomalies in large volume of data and anticipate and predict outcome

The Process of Building a Cognitive Application.

- Defining the objective
- defining the domain
- Understanding the intended users and their attributes
- defining question and exploring insights
- acquire the relevant data source
- Creating and refining the corpus
- training & Testing

References

- https://www.youtube.com/watch?v=GiyMGBuu45w&t=3s
- https://www.youtube.com/watch?v=iGmHnICXDss&t=12s
- https://www.slideshare.net/HansiThenuwara/natural-language-processing-64271235
- https://www.tutorialspoint.com/natural_language_processing/natural_language_processing_syntactic_analysis.htm
- https://www.tutorialspoint.com/natural_language_processing/natural_language_processing_syntactic_analysis.htm
- https://www.youtube.com/watch?v=DPi5rfTX6mw&t=1050s
- https://www.youtube.com/watch?v=SFQ-owZaU_s
- https://personal.utdallas.edu/~sanda/courses/NLP/Lecture10.pdf
- https://nlp.stanford.edu/projects/project-induction.shtml
- https://nlp.stanford.edu/IR-book/html/htmledition/types-of-language-models-1.html
- “Cognitive Computing and Big Data analytics” by Judith S. Hurwitz, Marcia Kaufman, Adrian Bowels ,

Thank You