

Keywords and Concepts for Test Review IAL 620

Linguistics (basic dictionary definition/understanding)

- Grammar
- Syntax
- Semantics
- Parts of Speech
- Named entity recognition
- Tokens and N-Grams
- Collocations
- Translation
- Ambiguity
- Polysemy

Pragmatics (basic dictionary definition/understanding)

- Context
- Prosody
- Discourse
- Code-switching
- Multimodal
- Speech Acts
- Dictionary
- Lexicon
- Corpus

Concepts from Readings

Early “problems” with NLP

- Context
- Discourse
- Polysemy
- Ambiguity

Bag of Words

- What is it?
- How does it avoid the early problems of NLP?
- Also described as “probabilistic” approach to language
- Studying “patterns of use” and frequency as opposed to “grammaticality” or language rules

History

Turing Test

- Limitations?
- Advantages?

ARPA/Internet

- What is the definition of “real time”?
- What are the names of the first “bots” to have an “all computer chat,” or bot-to-bot conversation, without humans?
- Who is credited with developing the @ sign for email?
- What is the relevance of MsgGroup?
 - Free Speech
 - Emoji creation
 - First virtual community
- What year was TCP/IP adopted?
- What organization created the World Wide Web and the first internet browser?

Web Scraping Basics

Be able to identify the following and their basic function:

- HTML tag
 - <p>
 - <a>
 - <h1> to <h6>
 - <time>
 -
- HTML attribute
 - href
 - datetime
 - alt
 - title
 - src
 - id
- CSS class
- Xpath

Best practices:

- robots.txt