# Concept Vocab List

Knowledge graph embedding

Natural language processing

Entity linking

Named entity recognition

Relation extraction

Graph neural networks

Word embeddings

Semantic representation

Ontology learning

Information extraction

Natural Language Processing

NLP

machine translation

information extraction

structured data

database

text summarization

News Blaster

multi-document

dialogue systems

understanding

generation

linguistics

probability

statistics

computer science

Ambiguity

Interpretation

Parsing

Syntactic structures

Word sense ambiguity

Anaphora

Pronoun

Discourse level

Speech recognition.

language modeling

statistical natural language processing

trigram language models

evaluation of language models

estimation techniques

linear interpolation

discounting methods

vocabulary (set V)

set of possible sentences

well-formed sentences

STOP symbol

training sample

distribution over sentences

probability of sentences

web data

billion words

New York Times

example sentences

sentence formation

language of interest

effectiveness of language models

natural language applications.

speech recognition

part of speech tagging

natural language pausing

acoustic input

Fourier analysis

training corpus

NLP

Markov Processes

joint probability distribution

vocabulary

chain rule of probabilities

conditional probability

sequence modeling

first order Markov Process

product of expressions

random variables

vocabulary size

fixed length sequences

language generation

natural language understanding

text processing

Natural Language Processing,

Markov processes,

first-order Markov,

second-order Markov,

third-order Markov,

joint probability,

Distribution,

random variables,

sequence,

Conditioning,

language modeling,

Vocabulary,

STOP symbol,

sampling process.

Language modeling

Trigram language models

Markov processes

Vocabulary

Parameters

Probability distribution

Maximum likelihood estimate

Training set

Counts

Estimation problem.

Language model

OpenAI

GPT-3.5 architecture

Trained

Knowledge cutoff

Text generation

Language understanding

Sentiment analysis

Part-of-speech tagging

Machine learning

Neural networks

Deep learning

Computational linguistics

language modelling

perplexity

maximum likelihood estimate

sparse data problem

bigram estimate

unigram estimate

bias-variance trade-off

context

Knowledge Graph

Embedding

Natural Language Processing

Named Entity Recognition

Entity Linking

Semantic Role Labeling

Word Embedding

Relation Extraction

Graph Convolutional Networks

Entity Alignment

Entity recognition

Named entity recognition (NER)

Part-of-speech (POS) tagging

Dependency parsing

Coreference resolution

Semantic role labeling (SRL)

Word sense disambiguation (WSD)

Knowledge Graph Embedding

Entity Recognition

Text Representation

Graph Embedding

Text Mining

Semantic Analysis

Information Extraction

Embeddings

Named Entity Recognition (NER)

Part-of-Speech (POS) Tagging

Word Sense Disambiguation (WSD)

Graph Neural Networks (GNNs)

Deep Learning.

List

Entity

Relation

Embedding Learning

Graph Neural Networks

Semantic Similarity

Word Embeddings

Dependency Parsing

Knowledge Graph Construction

Knowledge Graph Completion

Knowledge Graph Representation Learning

Natural Language Understanding

Text Generation

NLP Pipeline

Content Selection

Lexical Choice

Sentence Structure Generation

Discourse Structure

FOG system

Weather Forecast Reports

Columbia University

Bellcore

FUF and SURGE systems

Choices in Content, Coherence, Style, and Media

Syntactic Structure

Aggregating Facts

Referring Expressions

Statistical machine translation

IBM models

Tree-to-tree translations

Phrased-based machine translation

Syntax-based machine translation

Discriminative reranking

Syntactic structure

Synchronous grammar

Evaluation of machine translation

Adequacy

Grammaticality

BLEU (Bilingual Evaluation Understudy) metric

Unigram, bigram, trigram, fullgram overlap

Edit cost

Noisy channel methods

Modern machine translation

Encoder

Decoder

Translation model

Bayesian theorem

Trigram

Text to text generation

Summarization

Text to signal

Speech recognition

OCR

Spelling correction

IBM model

Linguistic achievement

Engineering achievement

Target language

Source language.

Machine translation

Warren Weaver

Information theory

Google Translate

Recipe

Translation issues

Syntactic correctness

Word order

Subject-verb-object

Verb-subject-object

Inflection

Lexical distinctions

Adjective-noun order

Feminine/masculine agreement

Noun-verb agreement

Languages

Internet users

Translation systems

Rosetta Stone

Arcturan

Centauri

English

Spanish

Kevin Knight

Intergalactic language

Pattern recognition

Alignment

Garcia

Voon

Translation works.

Dialogue systems

Turn-taking

Conversational impliciture

Grice's maxims

Speech acts

Abbott and Costello skit

Who's on first

Baseball players

Implicature

Informative contribution

Ambiguity avoidance

Card game

Robot named GGG

Relevance maxim

Manner maxim

Quantity maxim

Quality maxim

Discourse analysis

Anaphoric expressions

Coherence

Examples of coherence in text

Rhetorical Structure Theory (RST)

Nucleus and satellite in RST

Referring expressions

Antecedent

Coreference

Named entity

MUC (MUC-7) competition

NIST (National Institute of Standards and Technology)

Automatic systems

Classification

Features

Knowledge graph embeddings

Text processing

Natural language processing (NLP)

Document analysis

Sentence analysis

Pronoun resolution

Entity resolution

Feature extraction.