Natural Language Processing for MIR

Sergio Oramas, Luis Espinosa-Anke, Shuo Zhang, Horacio Saggion, Xavier Serra

http://mtg.upf.edu/nlp-tutorial

Sergio Oramas
MTG, Universitat Pompeu Fabra
Barcelona, Spain

Luis Espinosa-Anke
TALN, Universitat Pompeu Fabra
Barcelona, Spain

Shuo Zhang
Georgetown University
Washington D.C., USA

Horacio Saggion
TALN, Universitat Pompeu Fabra
Barcelona, Spain

Xavier Serra
MTG, Universitat Pompeu Fabra
Barcelona, Spain

EXCELENCIA
MARÍA
DE MAEZTU

upf. Music Technology Group

taln research group

Objectives

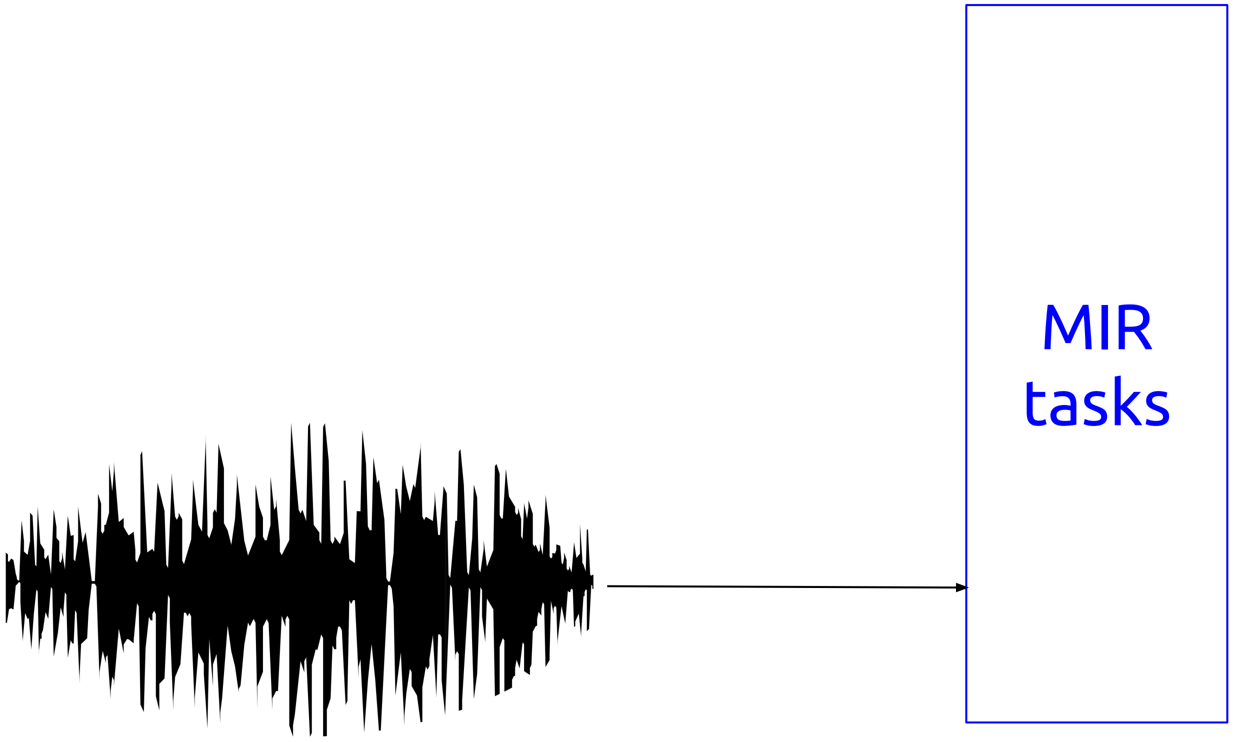
Provide a general introduction to NLP.

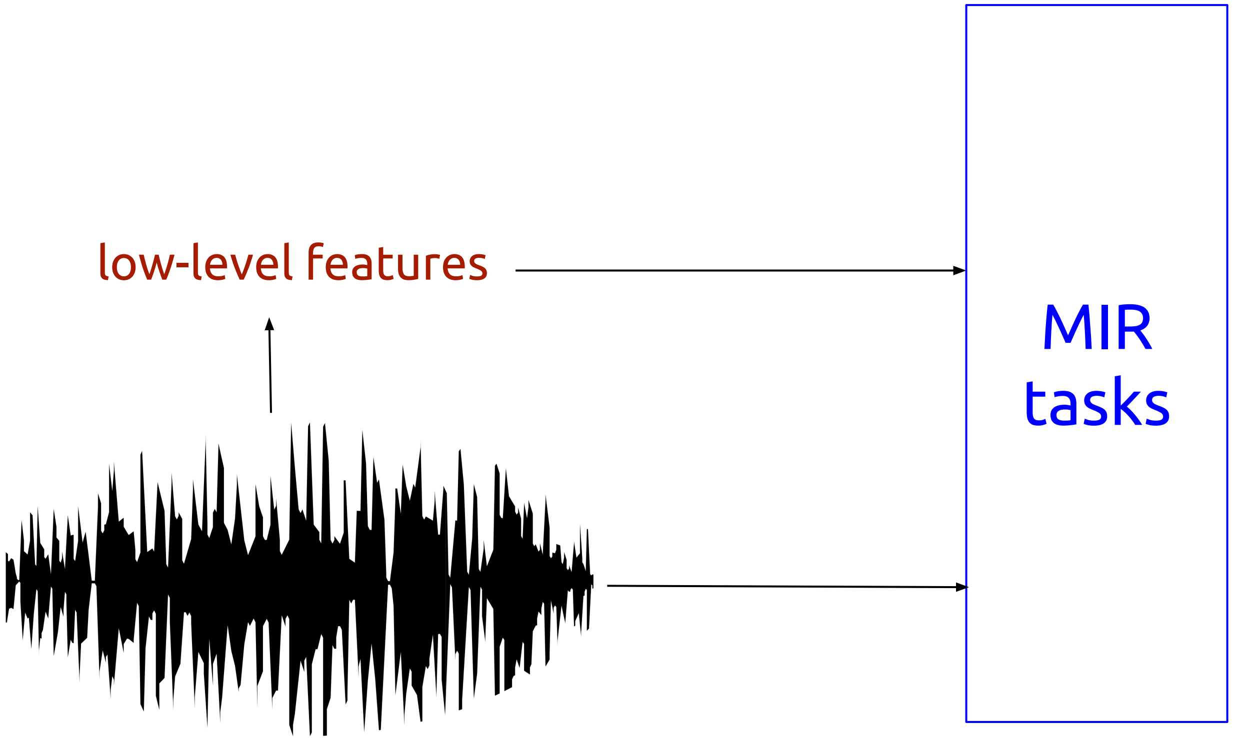
Identify areas of NLP with potential application in MIR.

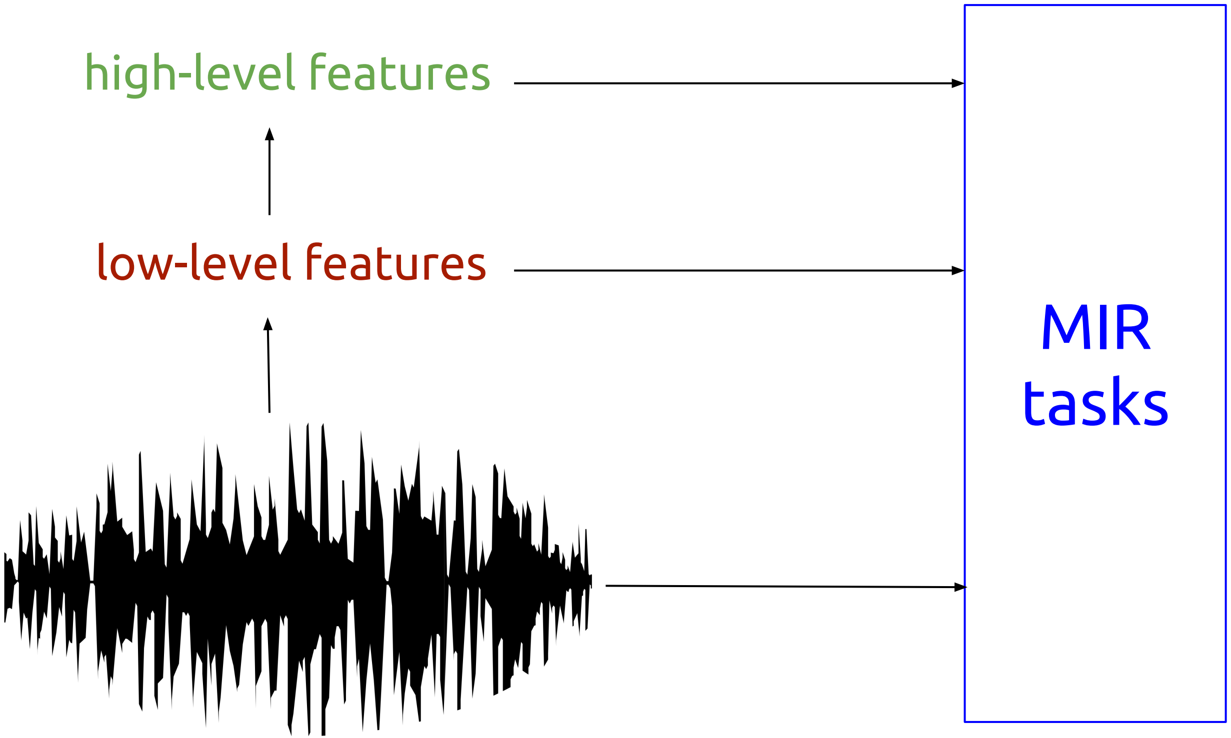
Address the extraction of semantic information from music text corpora.

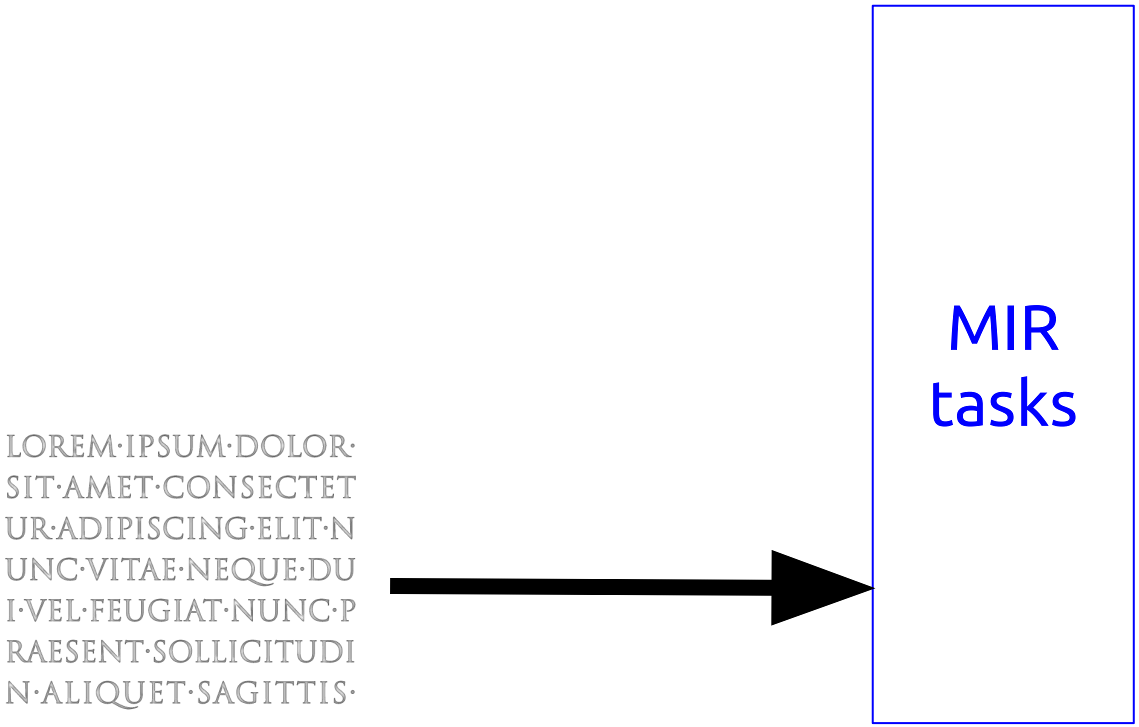
Show methodologies for exploiting semantic information in MIR.

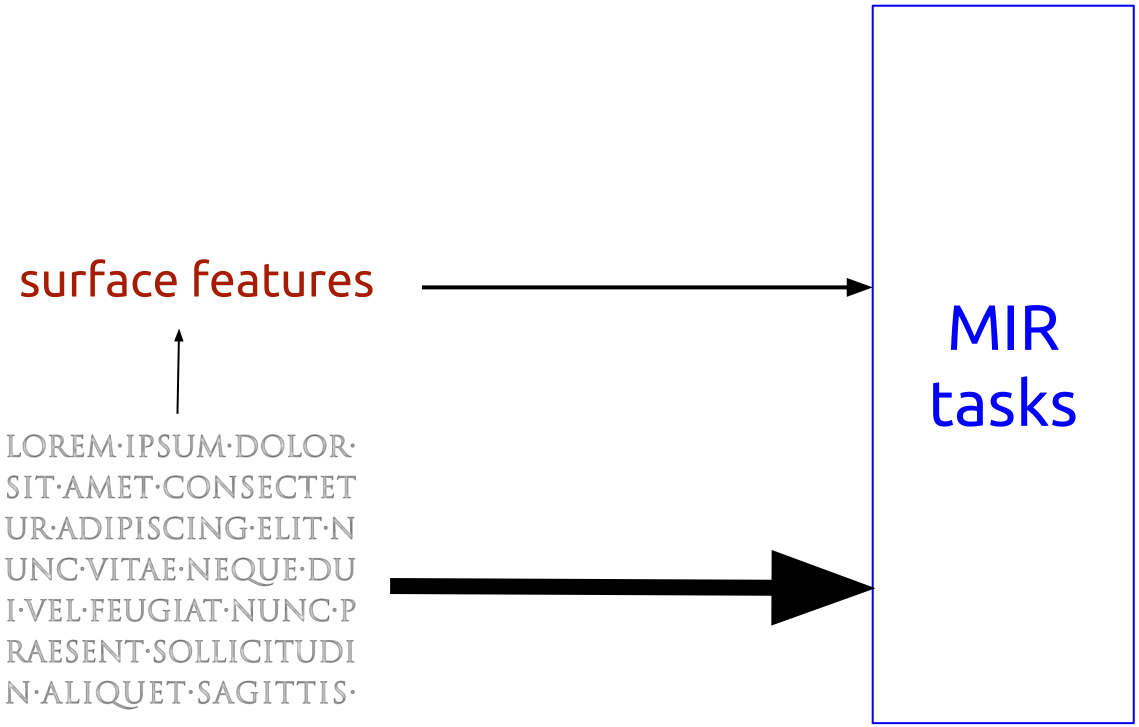
Why semantic information?

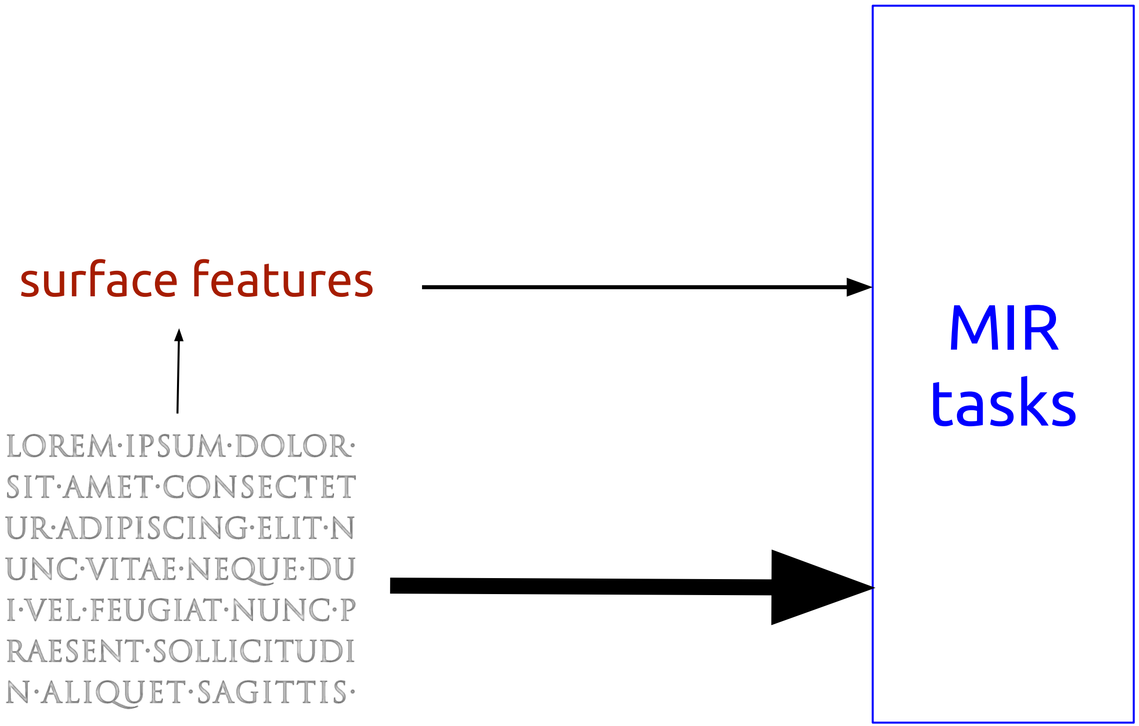

MIR
tasks


low-level features
MIR
tasks

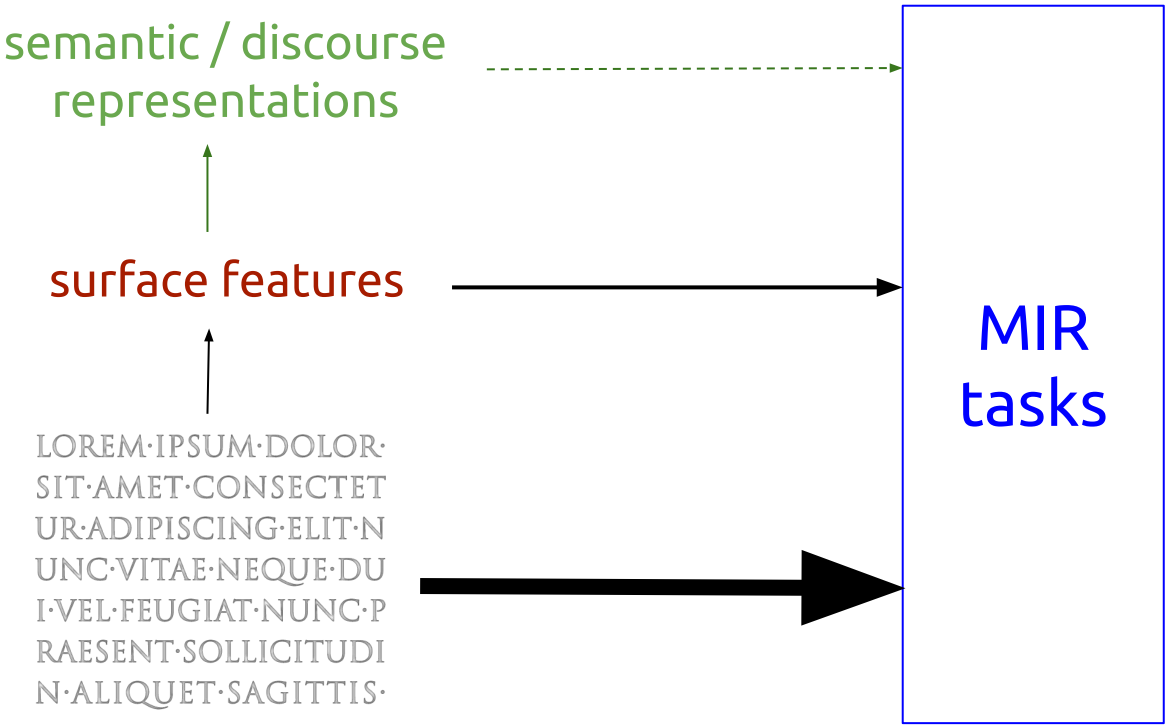

high-level features
low-level features
MIR
tasks


LOREM·IPSUM·DOLOR·
SIT·AMET·CONSECTET
UR·ADIPISCING·ELIT·N
UNC·VITAE·NEQUE·DU
I·VEL·FEUGIAT·NUNC·P
RAESENT·SOLLICITUDI
N·ALIQUET·SAGITTIS·
MIR
tasks


surface features
LOREM·IPSUM·DOLOR·
SIT·AMET·CONSECTET
UR·ADIPISCING·ELIT·N
UNC·VITAE·NEQUE·DU
I·VEL·FEUGIAT·NUNC·P
RAESENT·SOLLICITUDI
N·ALIQUET·SAGITTIS·
MIR
tasks

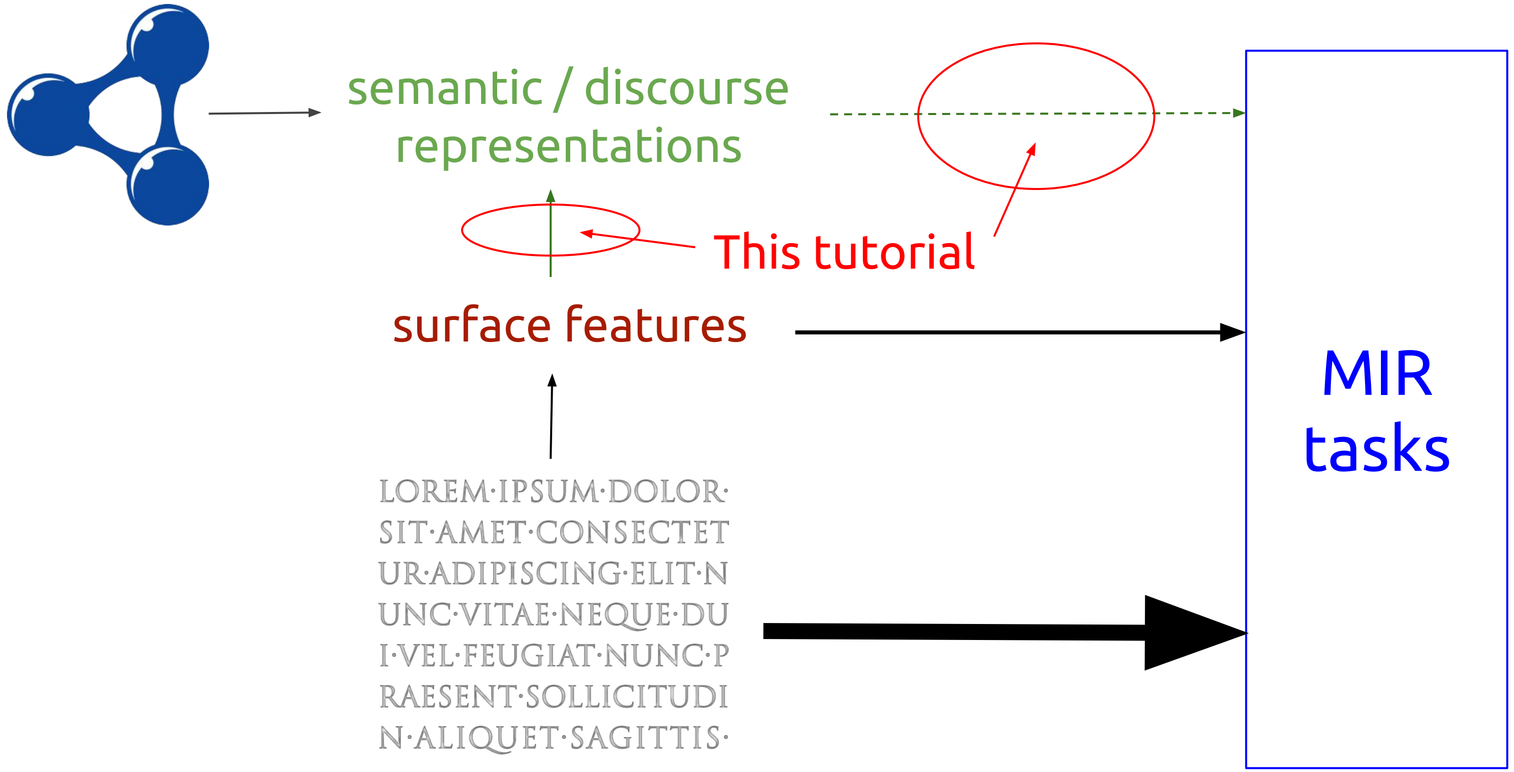


Peter Knees & Markus Schedl (2013): A Survey of Music Similarity and Recommendation from Music Context Data. ACM-TOMM.


semantic / discourse representations
surface features
LOREM·IPSUM·DOLOR·
SIT·AMET·CONSECTET
UR·ADIPISCING·ELIT·N
UNC·VITAE·NEQUE·DU
I·VEL·FEUGIAT·NUNC·P
RAESENT·SOLLICITUDI
N·ALIQUET·SAGITTIS·
MIR tasks

semantic / discourse representations

MIR tasks


semantic / discourse representations
This tutorial
surface features
LOREM·IPSUM·DOLOR·
SIT·AMET·CONSECTET
UR·ADIPISCING·ELIT·N
UNC·VITAE·NEQUE·DU
I·VEL·FEUGIAT·NUNC·P
RAESENT·SOLLICITUDI
N·ALIQUET·SAGITTIS·
MIR tasks

Text sources in MIR research

Lyrics

Biographies,
blogs, forums,
encyclopedias,
digital libraries,
social networks

Outline

- Introduction to NLP (20 mins)
- Information Extraction (15 mins)
  - Construction of Music Knowledge Bases (15 mins)
  - Applications in MIR (30 mins)

--- Coffee break ---

- Topic Modeling (20 mins)
- Sentiment Analysis (15 mins)
- Lexical Semantics (20 mins)
- Discussion (10 mins)

Outline

- **Introduction to NLP**
- Information Extraction
  - Construction of Music Knowledge Bases
  - Applications in MIR
- Topic Modeling
- Sentiment Analysis
- Lexical Semantics

Introduction to NLP

Outline

- What is Natural Language Processing?
- NLP Core Tasks
- Applications
- Knowledge Repositories
- Resources

What is Natural Language Processing?

· NLP is a field of Computer Science and Artificial Intelligence concerned with the interaction between computers and human (natural) language.

· Alan Turing's paper *Computing Machinery and Intelligence* is believed to be the first NLP paper. It stated that a computer could be considered intelligent if it could carry on a conversation with a human being without the human realizing he/she were talking to a machine.

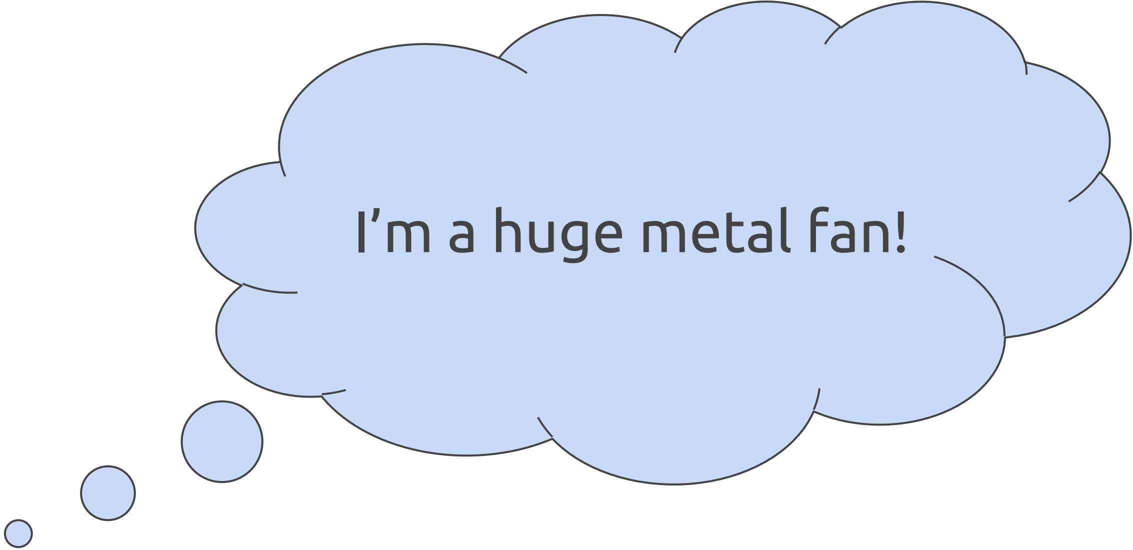
What is Natural Language Processing?

- There are over 7k languages in the world. Cultural and sociological traces

- “In the future, the most useful data will be the kind that was too **unstructured** to be used in the past.” [“The future of big data is quasi-unstructured,” Chewy Chunks, 23 March 2013] (from Wired.com).

- NLP is a core component in daily life technologies: web search, speech recognition and synthesis, automatic summaries in the web, product (including music) recommendation, machine translation...

Why is it hard?



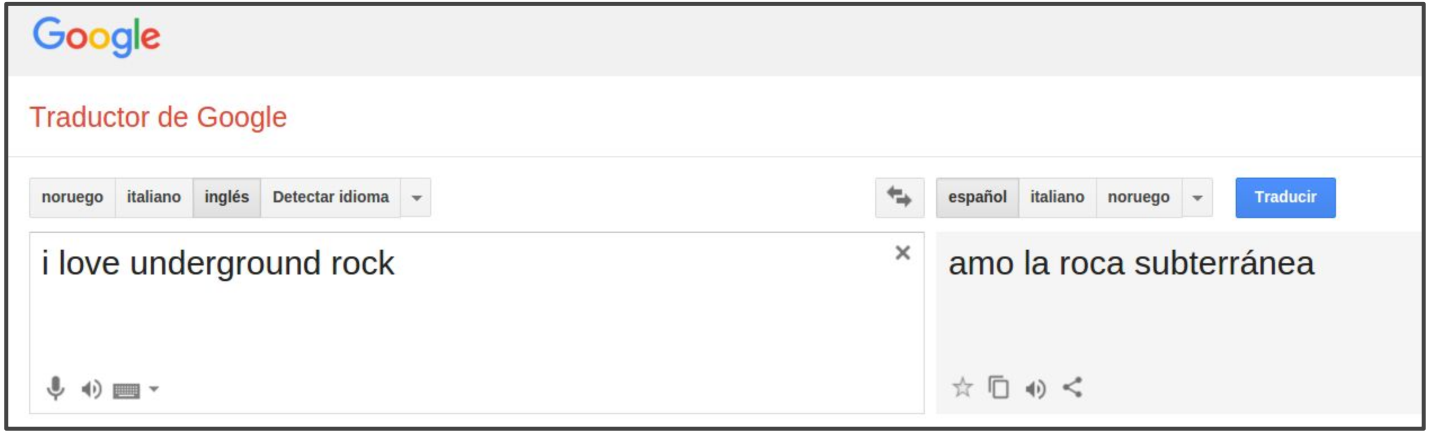
Why is it hard?

I'm a huge metal fan!

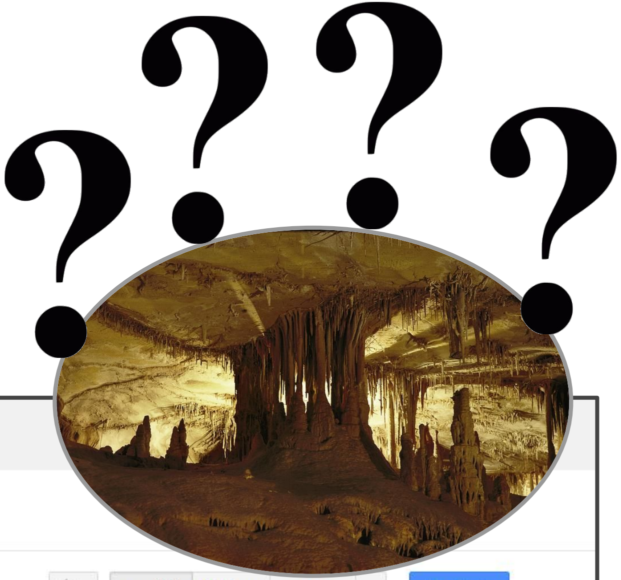
Why is it hard?

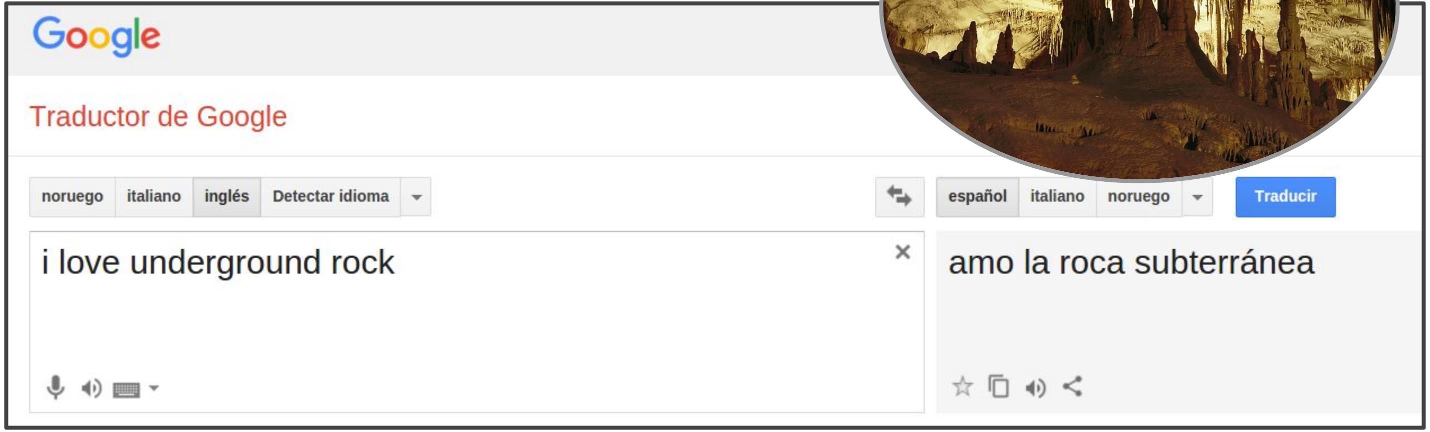
I'm a huge metal fan!

Why is it hard?



Why is it hard?



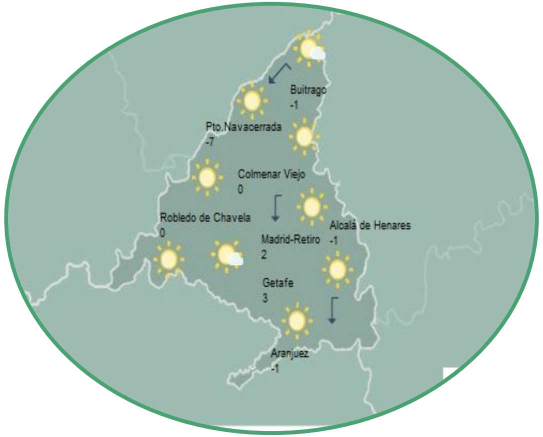


Why is it hard?

- “Plácido Domingo en Madrid”.

Why is it hard?

- “Plácido Domingo en Madrid”.



NLP is not a large uniform task

- **Core NLP Tasks**

    * Part-of-speech Tagging

    * Syntactic Parsing

    * Semantic Parsing

    * Named Entity Recognition

    * Coreference Resolution

    * Word Sense Disambiguation (WSD) & Entity Linking (EL)

Core elements in NLP - Part-of-Speech Tagging

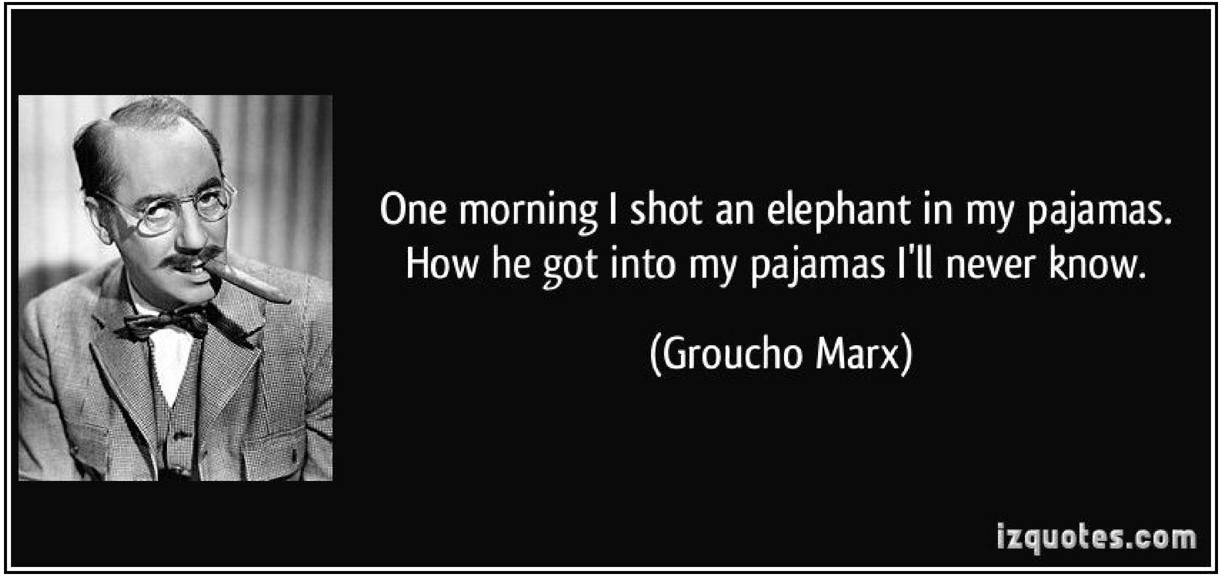
I like jazz music, it’s like being alive for a second.

Core elements in NLP - Part-of-Speech Tagging

I like jazz music, it’s like being alive for a second.

NOUN VERB NOUN NOUN PUNCT NOUN VERB ADP VERB ADJ ADP DET ADJ PUNCT

Core elements in NLP



http://www.nltk.org/book/ch08.html

Core elements in NLP - Syntactic Parsing

- Identify relations holding between words or phrases in the sentence, and what is their *function*.

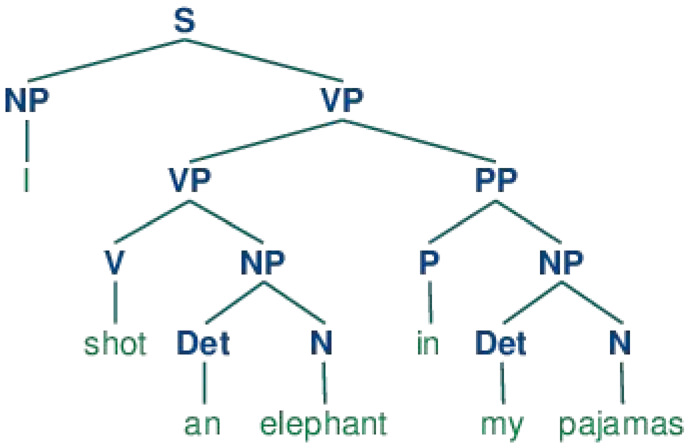
- By analyzing sentence structure, we understand the underlying meaning in a sentence.

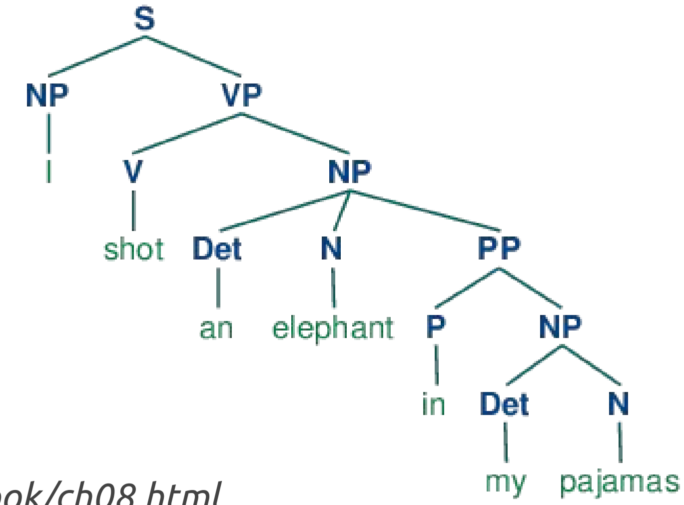
http://www.nltk.org/book/ch08.html

Core elements in NLP - Constituency Parsing

- Identify relations holding between words or phrases in the sentence, and what is their *function*.

- By analyzing sentence structure, we understand the underlying meaning in a sentence.



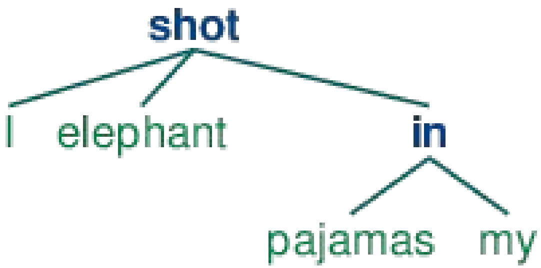


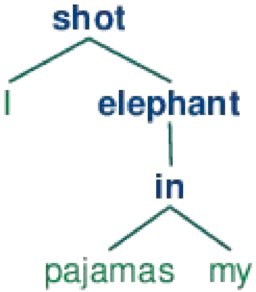
http://www.nltk.org/book/ch08.html

Core elements in NLP - Dependency Parsing

- Identify relations holding between words or phrases in the sentence, and what is their *function*.

- By analyzing sentence structure, we understand the underlying meaning in a sentence.





http://www.nltk.org/book/ch08.html

Core elements in NLP - Semantic Parsing

· A level of parsing above morphology and syntax. Capture underlying semantics expressed in language. Most focus on verbs and their *arguments*.

· A PropBank (http://propbank.github.io/) Example:

-> Mary *left* the room

* Arg0: **Entity leaving**, Arg1: **Place left**

-> Mary *left* her daughter her pearls

* Arg0: **Giver**, Arg1: **Thing given**, Arg2: **Beneficiary.**

Core elements in NLP - Named Entity Recognition

· Manfred Mann's Earth Band is a British progressive rock group formed in 1971 by Manfred Mann, a South African born keyboard player best known as a founding member and namesake of 60s group Manfred Mann.

Band
Music Genre
Artist
Country

Core elements in NLP - Coreference Resolution

"I voted for Nader because he was most aligned with my values," she said.

Core elements in NLP - **WSD** and EL

- “The performance of that bass player was outstanding”

Core elements in NLP - **WSD** and EL

- "The performance of that bass player was outstanding"

https://tackyraccoons.com/2011/11/21/all-your-bass-are-belong-to-us/

NLP is not a large uniform task

- **NLP Tasks**

    * Summarization
    * Simplification
    * Author Profiling
    * Machine Translation
    * Topic Modeling
    * Information Extraction

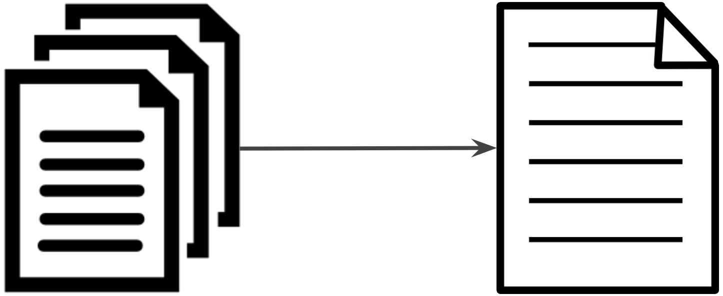
NLP Tasks - Summarization

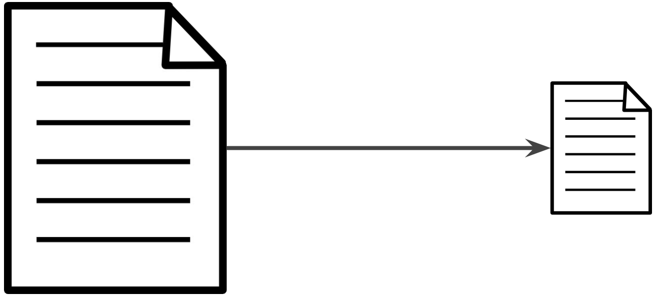
- Extractive

  * Retains most important sentences.

- Abstractive

  * Reformulates most important info.





NLP Tasks - Author Profiling

· Revealing demographic traces behind the writer of a message (*cybersecurity*), *aka* digital text forensics.

* From PAN 2016

```
<author id="{author-id}"
    lang="en|es|nl"
    age_group="18-24|25-34|35-49|50-64|65-xx"
    gender="male|female"
/>
```

NLP Tasks - Machine Translation

· Given text in L1, translate it into L2.

· One of the most widely known NLP tasks

· Originally it was approached as a rule-based task. Today, statistical approaches have taken over.

· Apertium is one of the best known RBMT systems (www.apertium.org).

· SMT is, by far, the most studied MT discipline. Challenges include *sentence alignment*, *word alignment*, *statistical anomalies*, *idioms*, *different word orders*, *OOV*.

Knowledge Repositories and Knowledge Bases

· A Knowledge Base (KB) is a rich form of Knowledge Repository (KR), term coined to differentiate from traditional *databases*.

· The term KB may be used to refer to terminological or lexical databases, ontologies, and any graph-like KR.

· KBs are essential for AI tasks such as reasoning, inference or semantic search. Also for Word Sense Disambiguation, Entity Linking, Machine Translation, Semantics…

· They may be constructed manually in specific domains (e.g. *Chemistry*), but the general preference is to learn them (semi) automatically.

Knowledge Bases

- Hand-crafted KBs

  → From generic to domain-specific. E.g. **WordNet**, **CheBi**, **SnomedCT**.

- Integrative Projects

  → Unify in one single resource manually curated KRs and KBs.

  ⇒ **BabelNet** (originally, WordNet + Wikipedia), **DBPedia**, **Yago**…

- Open Information Extraction for KB construction

  → **NELL**, **PATTY**, **WiseNet**, **DefIE**, **KB-Unify**...

Music Knowledge Bases

- **MusicBrainz** and **Discogs**
    - → Open encyclopedias of music metadata
    - → MB is regularly published as Linked Data by the LinkedBrainz project.
- **Grove Music Online**
    - → Music *scholar* encyclopedia
- **Flamenco MKB**

Software

Standalone

· OpenNLP: https://opennlp.apache.org/

· Stanford CoreNLP: http://stanfordnlp.github.io/CoreNLP/

· Freeling: http://nlp.lsi.upc.edu/freeling/node/1

· Gate: https://gate.ac.uk/

· Mate Parser:
http://www.ims.uni-stuttgart.de/forschung/ressourcen/werkzeuge/matetools.en.html

Python Libraries

· Spacy: https://spacy.io

· Pattern: http://www.clips.ua.ac.be/pattern

· NLTK: http://www.nltk.org/

· Gensim: https://radimrehurek.com/gensim/

· Blob: http://textblob.readthedocs.io/en/dev/

· Rake:
https://www.airpair.com/nlp/keyword-extraction-tutorial

Software

ML toolkits/libraries widely used in NLP

· CRF++: https://taku910.github.io/crfpp/

· Mallet: http://mallet.cs.umass.edu/

· Networkx: https://networkx.github.io

· Weka: http://www.cs.waikato.ac.nz/ml/weka/

References - NLP

Part-of-Speech Tagging: Schmid, H. (1994, September). Probabilistic part-of-speech tagging using decision trees. In Proceedings of the international conference on new methods in language processing (Vol. 12, pp. 44-49).

Parsing: Chomsky, N. (2002). Syntactic structures. Walter de Gruyter. ; Nivre, J. (2003). An efficient algorithm for projective dependency parsing. In Proceedings of the 8th International Workshop on Parsing Technologies (IWPT).

Named Entity Recognition: Tjong Kim Sang, E. F., & De Meulder, F. (2003, May). Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4 (pp. 142-147). Association for Computational Linguistics.

Coreference Resolution: Soon, W. M., Ng, H. T., & Lim, D. C. Y. (2001). A machine learning approach to coreference resolution of noun phrases. Computational linguistics, 27(4), 521-544.

Summarization: Saggion, H., & Lapalme, G. (2002). Generating indicative-informative summaries with sumUM. Computational linguistics, 28(4), 497-526.

Simplification: Chandrasekar, R., Doran, C., & Srinivas, B. (1996, August). Motivations and methods for text simplification. In Proceedings of the 16th conference on Computational linguistics-Volume 2 (pp. 1041-1044). Association for Computational Linguistics.

Sentiment Analysis: Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1-135.

Author Profiling: Estival, D., Gaustad, T., Pham, S. B., Radford, W., & Hutchinson, B. (2007). Author profiling for English emails. In Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING'07) (pp. 263-272).

Topic Modeling: Wallach, H. M. (2006, June). Topic modeling: beyond bag-of-words. In Proceedings of the 23rd international conference on Machine learning (pp. 977-984). ACM.

Machine Translation: Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., ... & Dyer, C. (2007, June). Moses: Open source toolkit for statistical machine translation. In Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions (pp. 177-180). Association for Computational Linguistics.

Lexical Semantics: Cruse, D. A. (1986). Lexical semantics. Cambridge University Press.

Word Sense Disambiguation. Navigli, R. (2009). "Word sense disambiguation: A survey." ACM Computing Surveys (CSUR) 41.2: 10.

References - KBs

WordNEt: Miller, George A. "WordNet: a lexical database for English." *Communications of the ACM* 38.11 (1995): 39-41.

Chebi: Degtyarenko, Kirill, et al. "ChEBI: a database and ontology for chemical entities of biological interest." Nucleic acids research 36.suppl 1 (2008): D344-D350.

Snomed: Spackman, Kent A., Keith E. Campbell, and Roger A. Côté. "SNOMED RT: a reference terminology for health care." Proceedings of the AMIA annual fall symposium. American Medical Informatics Association, 1997.

BabelNet: Navigli, Roberto, and Simone Paolo Ponzetto. "BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network." Artificial Intelligence 193 (2012): 217-250.

DBPedia: Auer, Sören, et al. "Dbpedia: A nucleus for a web of open data." The semantic web. Springer Berlin Heidelberg, 2007. 722-735.

Yago: Suchanek, Fabian M., Gjergji Kasneci, and Gerhard Weikum. "Yago: a core of semantic knowledge." Proceedings of the 16th international conference on World Wide Web. ACM, 2007.

NELL: Carlson, Andrew, et al. "Toward an Architecture for Never-Ending Language Learning." AAAI. Vol. 5. 2010.

PATTY: Nakashole, Ndapandula, Gerhard Weikum, and Fabian Suchanek. "PATTY: a taxonomy of relational patterns with semantic types." *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Association for Computational Linguistics, 2012.

WiseNet: Moro, Andrea, and Roberto Navigli. "WiSeNet: Building a Wikipedia-based semantic network with ontologized relations." Proceedings of the 21st ACM international conference on Information and knowledge management. ACM, 2012.

DefIE: Delli Bovi, Claudio, Luca Telesca, and Roberto Navigli. "Large-Scale Information Extraction from Textual Definitions through Deep Syntactic and Semantic Analysis." Transactions of the Association for Computational Linguistics 3 (2015): 529-543.

KB-Unify: Bovi, Claudio Delli, Luis Espinosa Anke, and Roberto Navigli. "Knowledge Base Unification via Sense Embeddings and Disambiguation." Proceedings of EMNLP. 2015.

MusicBrainz: Swartz, Aaron. "Musicbrainz: A semantic web service." IEEE Intelligent Systems 17.1 (2002): 76-77.

Discogs: www.discogs.com

Grove Online: http://www.oxfordmusiconline.com/public/book/omo_gmo

Several levels of *linguistic description*.

· **Phonetical:** Language sounds, phonemes different in *house* and *mouse*.

· **Morphological:** Morphemes and lexemes: *artist* -> *artists*, *sing* -> *sang*.

· **Syntactic:** Word combinations in higher structures.

I[SBJ] <- **love**[ROOT] -> **music**[DOBJ]

· **Semantic:** Lexical semantics (word meanings), textual semantics.

· **Textual:** Coherence and cohesiveness of texts. Coreference, theme development or intention.

· **Pragmatic:** Communicative contexts. World knowledge. User profiles.

Outline

- Introduction to NLP
- **Information Extraction**
    - Construction of Music Knowledge Bases
    - Applications in MIR
- Topic Modeling
- Sentiment Analysis
- Lexical Semantics

Information Extraction

Information Extraction

Information extraction (IE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents.

Unstructured vs. Structured

Information Extraction

Unstructured text

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .

Information Extraction

Entity Recognition

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .

Information Extraction

Entity Recognition and Classification

Organization

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .

Work of art

Person

Information Extraction

Wilco (disambiguation)

From Wikipedia, the free encyclopedia

Wilco is an American rock band.

Wilco may also refer to:

- Wilco (voice procedure), a radio procedure word, short for "**Will Co**mply"; origin of the term
- *Wilco (The Album)*, an album by the band Wilco, or the title song, "Wilco (The Song)"
- *Wilco: Learning How to Die*, a book about the band, by Greg Kot
- Wilco (farm supply cooperative), an American chain of agricultural cooperative stores
- Wilco (tree), *Anadenanthera colubrina*, a South American tree
- Wilkinson County, Georgia, sometimes abbreviated as "Wilco"
- Williamson County, Texas, sometimes abbreviated as "Wilco"
- WilcoHess, the chain of gas stations

Information Extraction

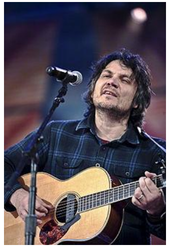
Entity disambiguation or Entity Linking

https://en.wikipedia.org/wiki/Wilco

WIKIPEDIA
The Free Encyclopedia

Organization

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .



Work of art

http://musicbrainz.org/recording/246500ae-379b-4290-8716-d58b596753dd

Person

https://en.wikipedia.org/wiki/Jeff_Tweedy

Hate It Here
~ Recording by Wilco

MusicBrainz

Information Extraction

Relation Extraction

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .

Hate It Here $\xrightarrow{\text{was written by}}$ Jeff Tweedy

Jeff Tweedy $\xrightarrow{\text{frontman}}$ Wilco

Information Extraction

Relation Extraction

Unstructured

"Hate It Here" was written by Wilco frontman , Jeff Tweedy .

Structured

Hate It Here $\xrightarrow{\text{was written by}}$ Jeff Tweedy

Jeff Tweedy $\xrightarrow{\text{frontman}}$ Wilco

Entity Linking

Entity linking is the task to associate, for a given candidate textual fragment, the most suitable entry in a reference Knowledge Base.

- Also referred to as Entity Disambiguation
- Typically Wikipedia, DBpedia, YAGO, Freebase as reference KB

Entity Linking

Entity linking is the task to associate, for a given candidate textual fragment, the most suitable entry in a reference Knowledge Base.

- Also referred to as Entity Disambiguation
- Typically Wikipedia, DBpedia, YAGO, Freebase as reference KB

Entity linking is typically broken down into **two main phases**:

- Candidate selection (entity annotation)
- Reference disambiguation (entity resolution)

Entity Linking

The entity linking system can either **return**:

- Matching entry (e.g. DBpedia URI, Wikipedia URL)
- NIL (no matching in the Knowledge Base)

But most of the systems make the closed world assumption, i.e. there is always a target entity in the knowledge base.

Entity Linking

Entity linking needs to handle:

- Name variations (entities are referred to in many different ways)
  - e.g. Elvis, Elvis Presley, Elvis Aaron Presley, The King of Rock and Roll
- Entity ambiguity (the same string can refer to more than one entity)
  - e.g. Prince, Debut, Bach, Strauss
- Missing entities (there is no target entity in the knowledge base)

Entity Linking: Tools

Babelfy: Entity Linking + Word Sense Disambiguation. Web service. KB: BabelNet. http://babelfy.org/index

Tagme: Web service. KB: Wikipedia. https://tagme.d4science.org/tagme/

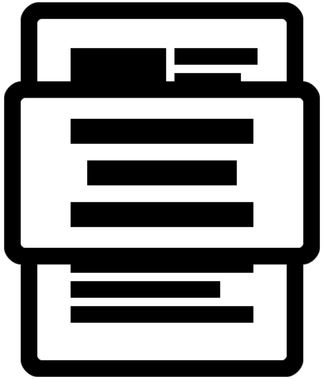
DBpedia Spotlight. Installable web service. KB: DBpedia. https://github.com/dbpedia-spotlight/dbpedia-spotlight

Relation Extraction

Detection and classification of semantic relations within a set of artifacts (e.g. entities, noun phrases) from text.

Numerous **variants**:

- Supervision: {fully, un, semi, distant}-supervision
- Undefined vs. pre-determined set of relations
- Binary vs. n-ary relations

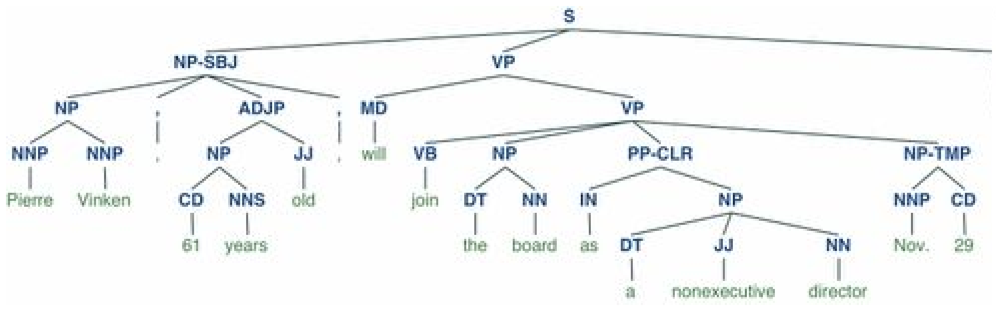


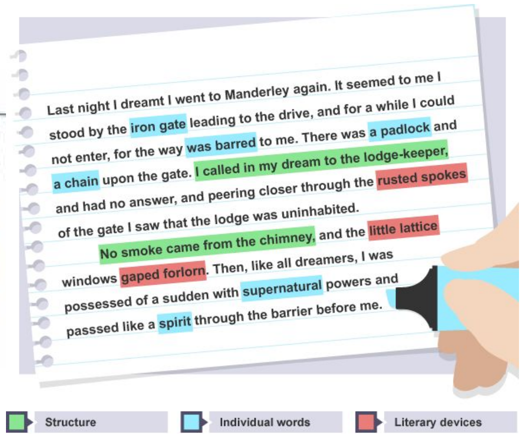
Created by Alex Getty
from Noun Project

Relation Extraction

Typical **features**:

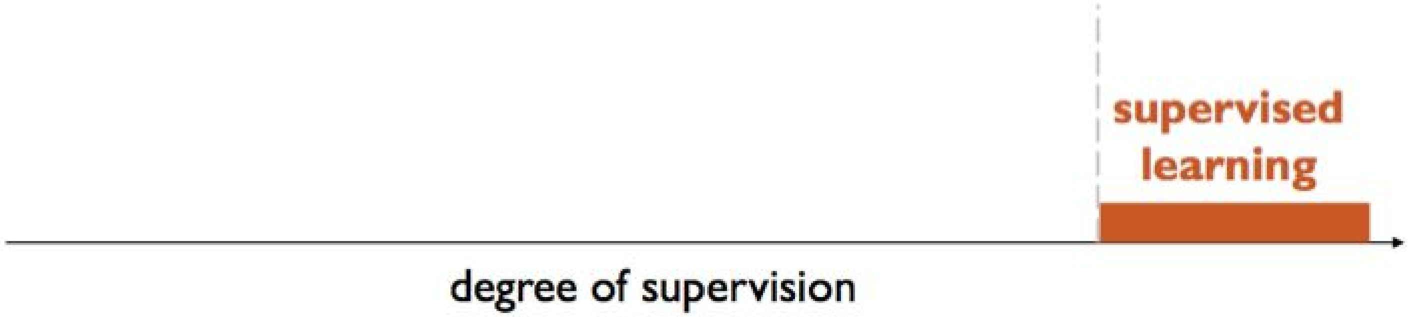
- morphologic, syntactic, semantic, statistical
- context words + part-of-speech tags, dependency paths, named entities





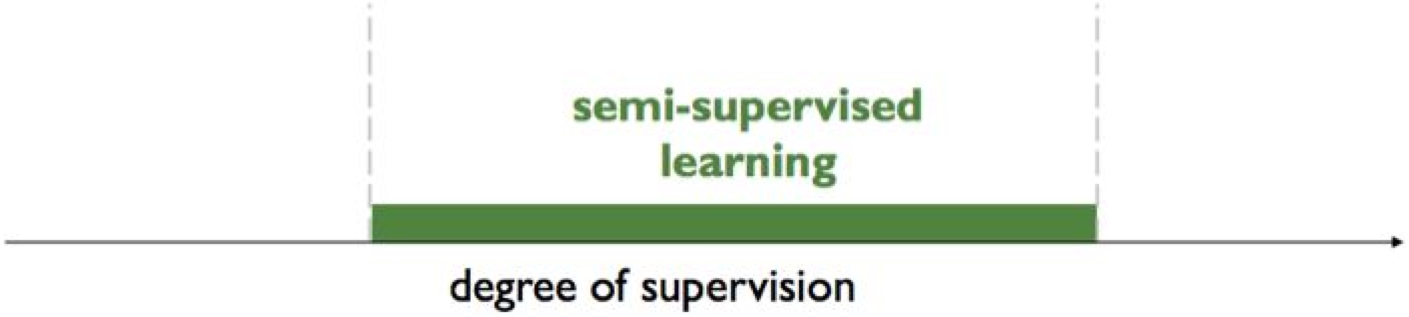
Relation Extraction

- Input:
  - Large corpus of unstructured text
  - Set of semantic relations
  - Labelled training data
- Output:
  - Knowledge Base of triples
  - 〈 entity, relation, entity 〉



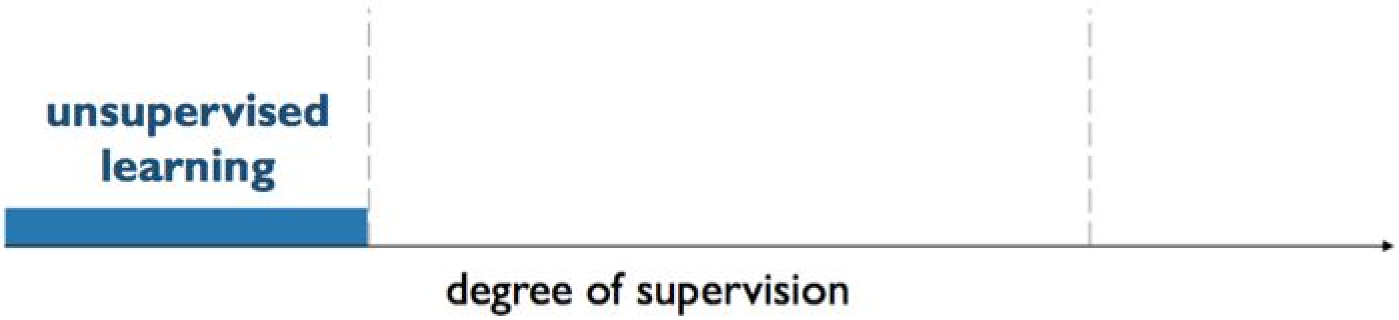
Relation Extraction

- Input:
  - Large corpus of unstructured text
  - Set of semantic relations
  - High-precision seeds/examples
- Output:
  - Knowledge Base of triples
  - 〈 entity, relation, entity 〉

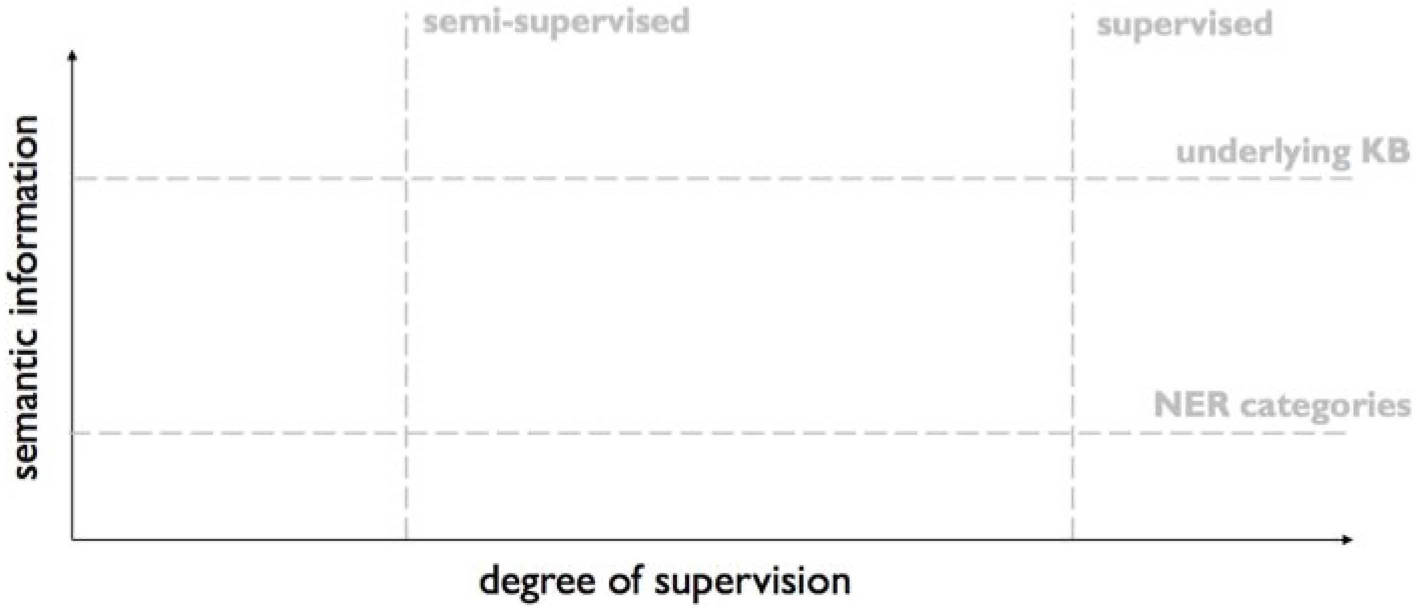


Relation Extraction

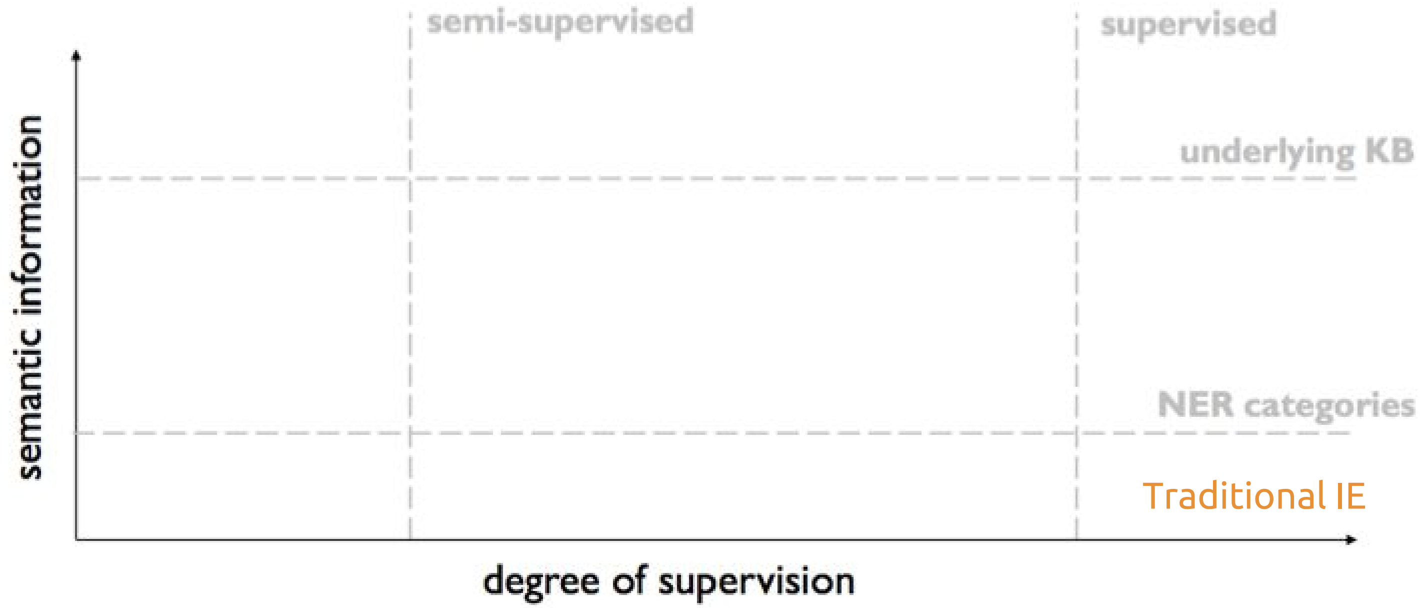
- Input:
  - Large corpus of unstructured text
  - ~~Set of semantic relations~~
  - ~~Labelled training data~~
- Output:
  - Knowledge Base of triples
  - 〈 entity, relation, entity 〉
  - Set of semantic relations



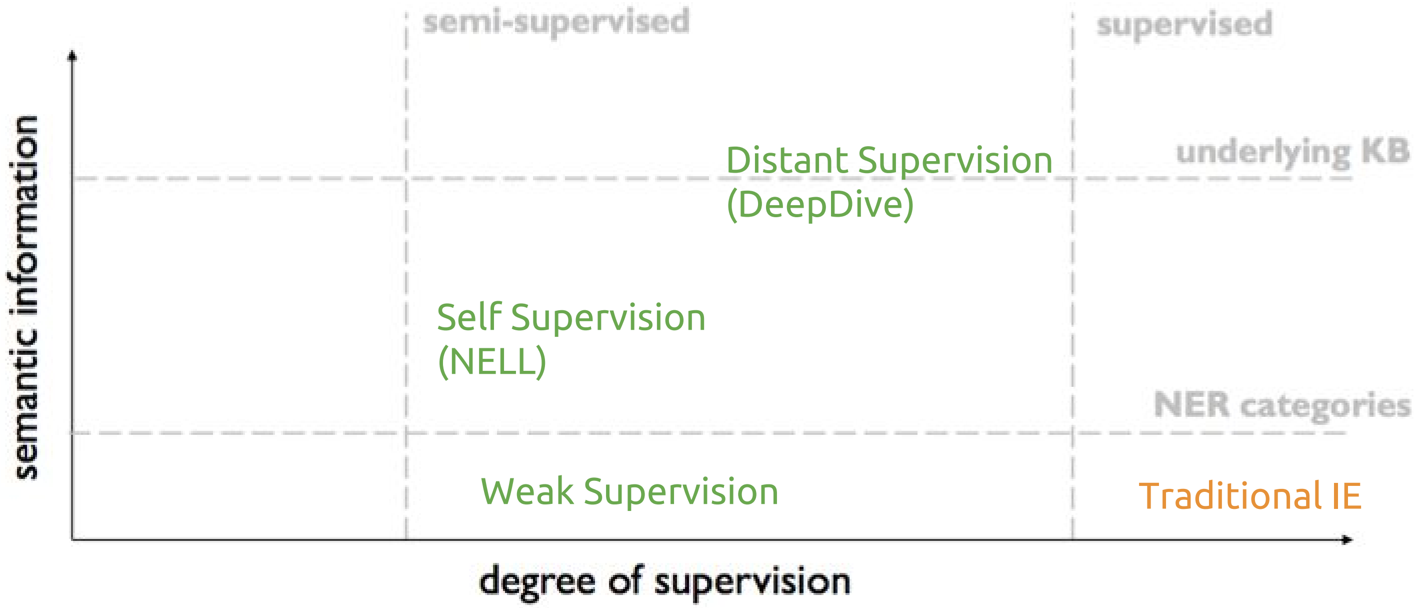
Relation Extraction



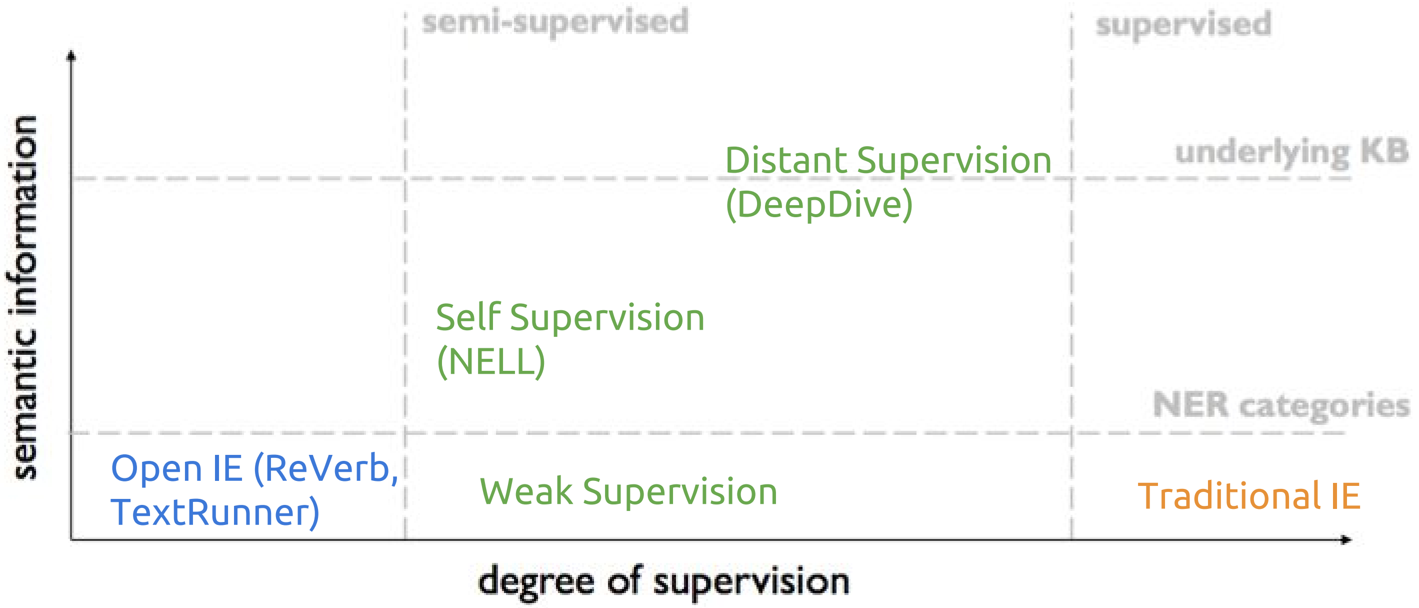
Relation Extraction



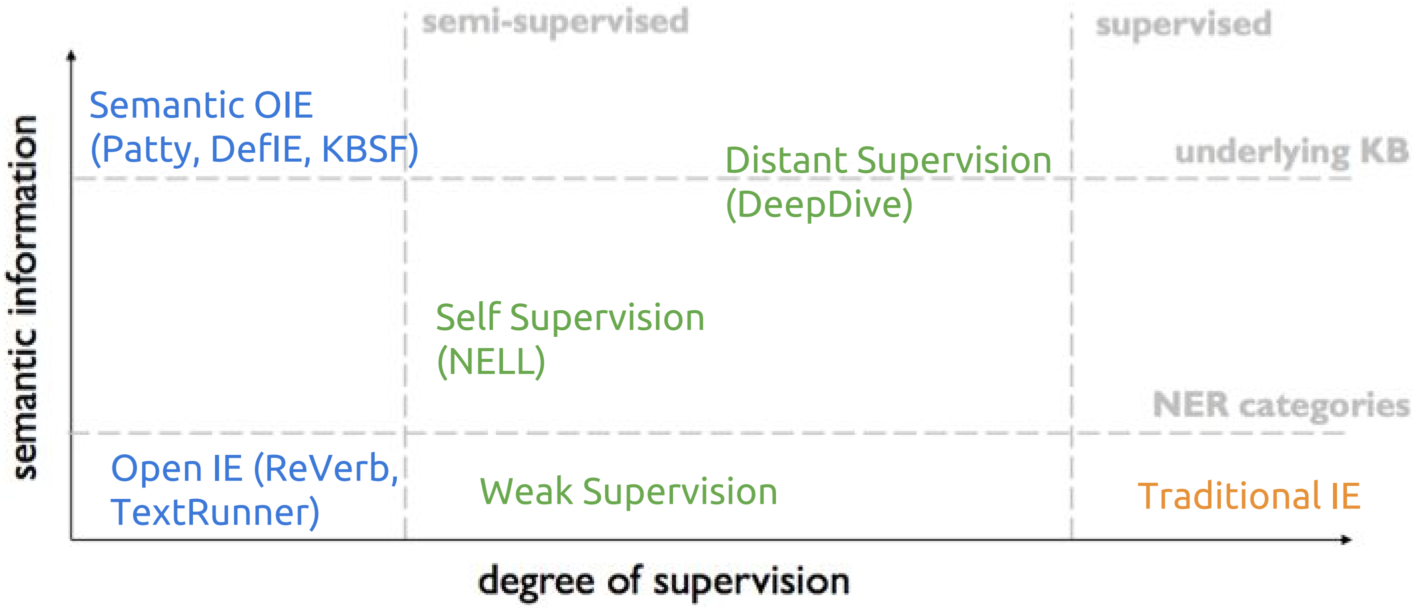
Relation Extraction



Relation Extraction



Relation Extraction



Semantic Open IE

Entity Linking + Open Information Extraction

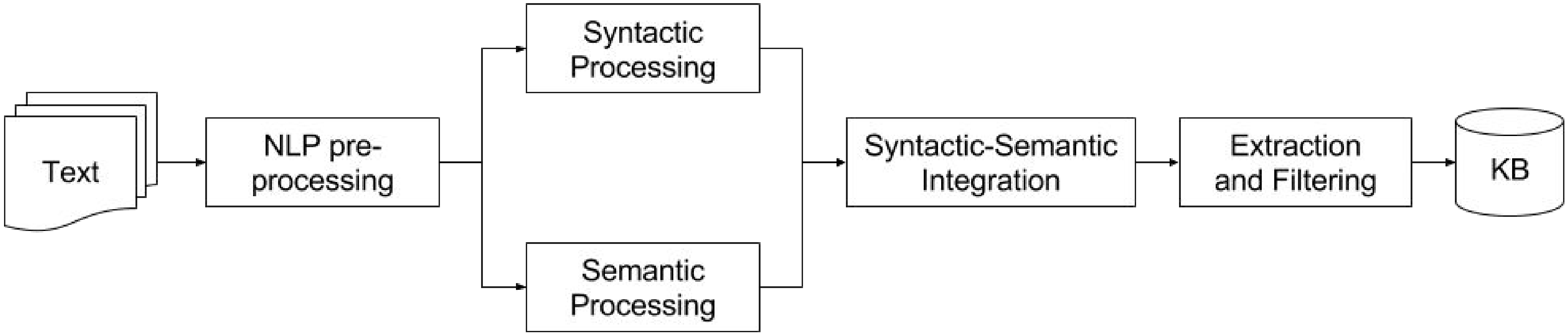
Advantages

- Not restricted to a set of predefined relations
- Unsupervised: no need of training samples
- Use of semantic information reduces imprecision of Open IE
- Useful for KB construction and KB expansion (no need of mapping)

Oramas S., Espinosa-Anke L., Sordo M., Saggion H., Serra X. (2016). *Information Extraction for Knowledge Base Construction in the Music Domain.* Journal on Knowledge & Data Engineering, Elsevier.

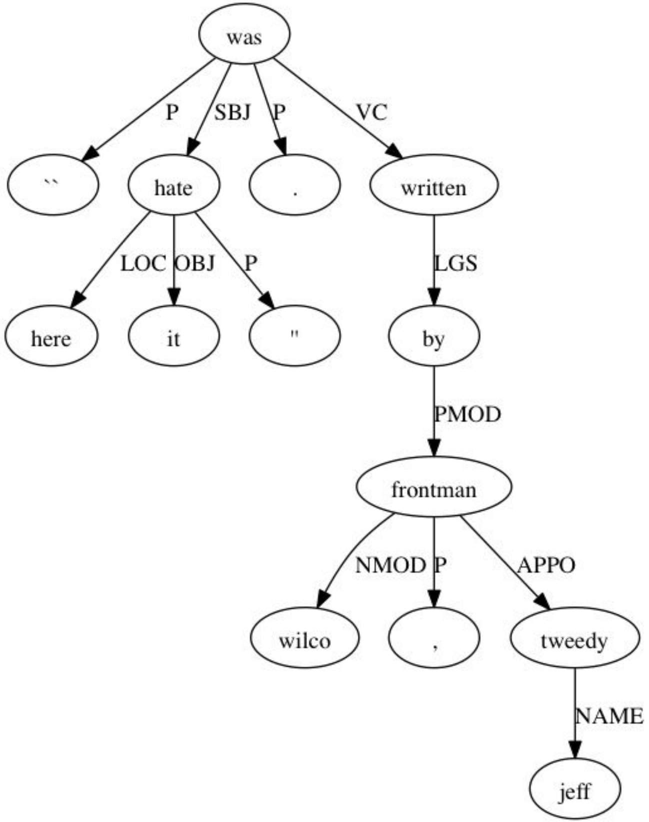
Semantic Open IE

- Entity linking -> Semantic Information
- Parsing (e.g. dependency parsing) -> Syntactic Information
- Semantic-Syntactic integration
- Shortest path between entities
- Filtering of relations



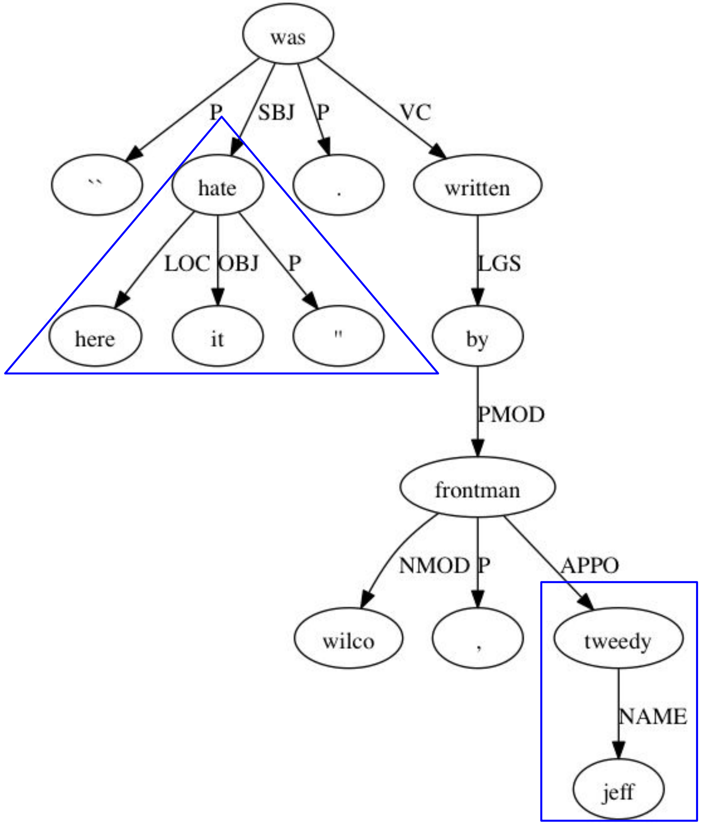
`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .

Semantic Open IE



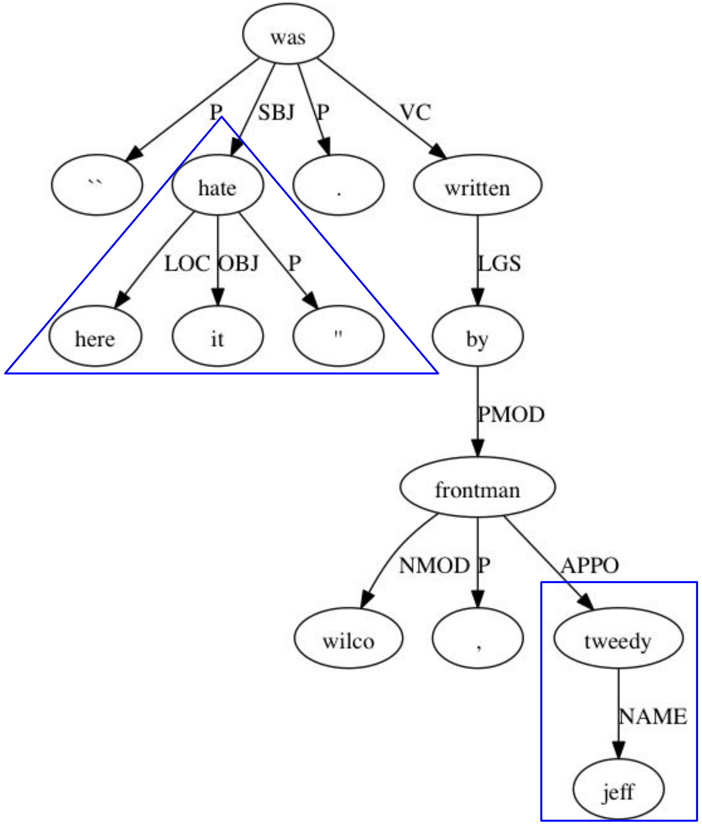
`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .

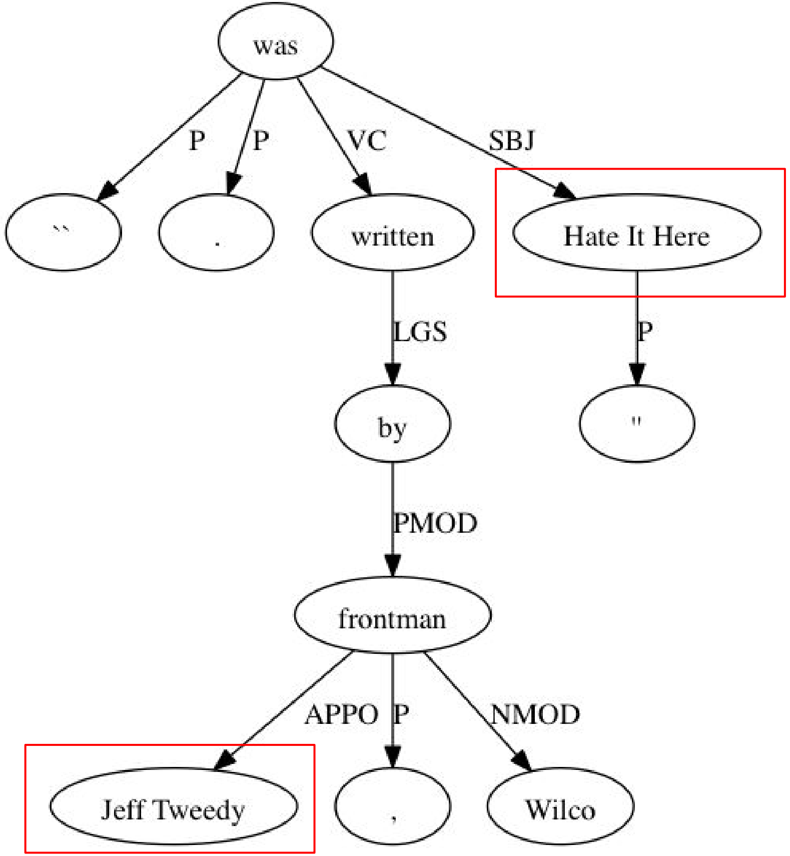
Semantic Open IE



`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .

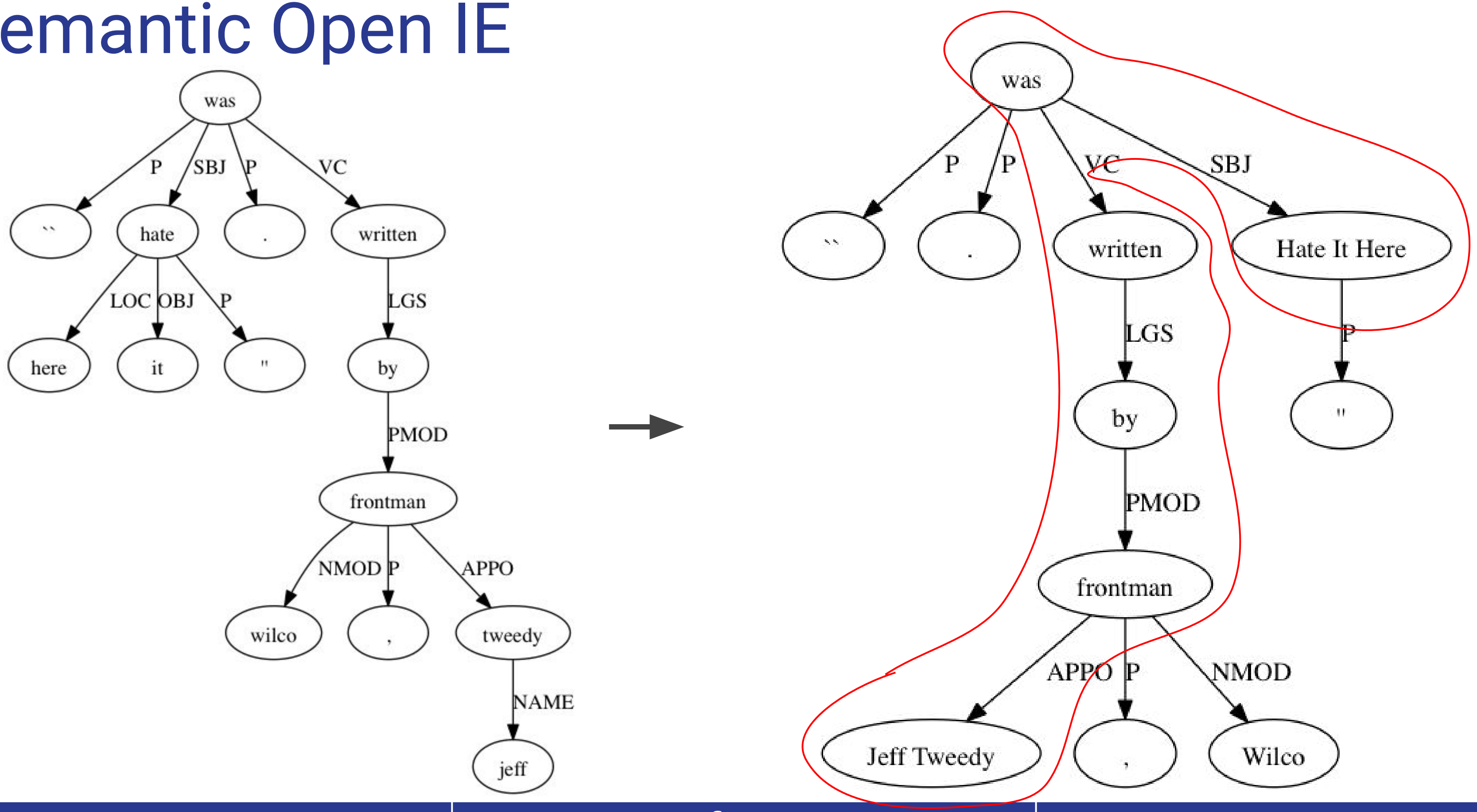
Semantic Open IE





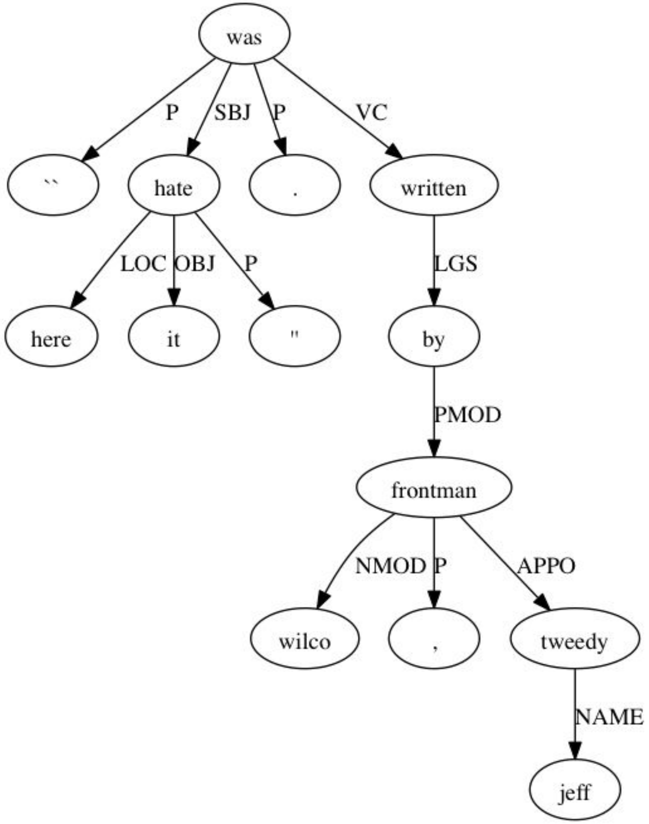
Semantic Open IE

`` Hate It Here " was written by Wilco frontman , Jeff Tweedy .

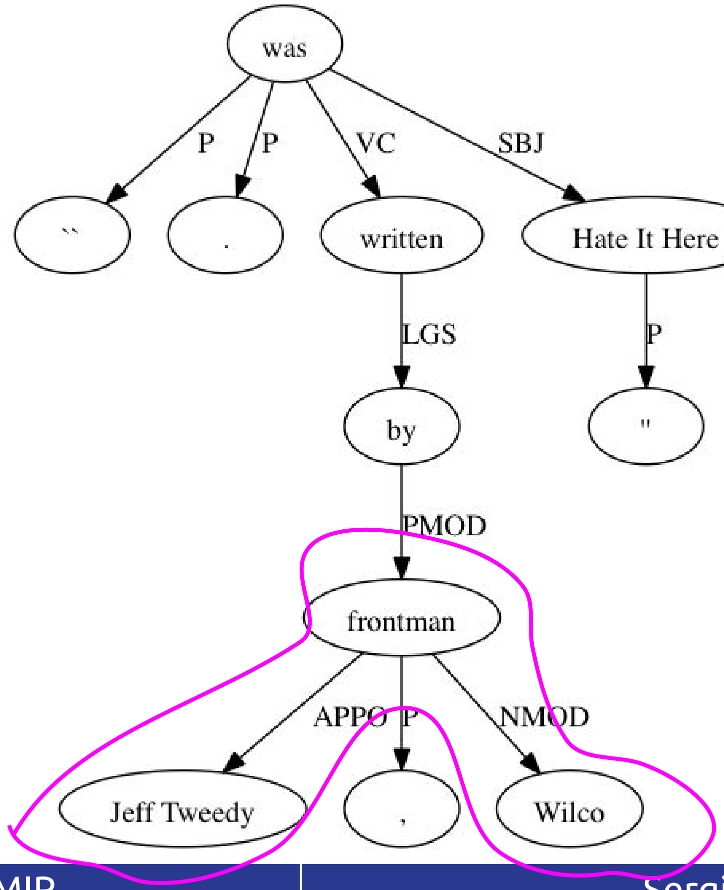


Semantic Open IE

`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .



→



Relation Extraction (References)

Traditional IE

Zhao, S., & Grishman, R. (2005). Extracting relations with integrated information using kernel methods. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics - ACL '05* (pp. 419–426).

Weak Supervision

Bunescu, R. C., & Mooney, R. J. (2007). Learning to Extract Relations from the Web using Minimal Supervision. *Computational Linguistics*, *45*(June), 576–583.

Self Supervision

Carlson, A., Betteridge, J., & Kisiel, B. (2010). Toward an Architecture for Never-Ending Language Learning. *In Proceedings of the Conference on Artificial Intelligence (AAAI) (2010)*

Distant Supervision

Riedel, S., Yao, L., & McCallum, A. (2010). Modeling relations and their mentions without labeled text. In *Lecture Notes in Computer Science* (Vol. 6323 LNAI, pp. 148–163).

Relation Extraction (References)

Open IE

Fader, A., Soderland, S., & Etzioni, O. (2011). Identifying relations for open information extraction. *Proceedings of the Conference on Empirical Methods in Natural Language Processing EMNLP '11*, 1535–1545.

Semantic Open IE

Nakashole, N., Weikum, G., & Suchanek, F. M. (2012). PATTY: A Taxonomy of Relational Patterns with Semantic Types. *EMNLP-CoNLL*, (July), 1135–1145.

Delli Bovi, C., Telesca, L., & Navigli, R. (2015). Large-Scale Information Extraction from Textual Definitions through Deep Syntactic and Semantic Analysis. *Transactions of the Association for Computational Linguistics*, *3*, 529–543.

Oramas S., Espinosa-Anke L., Sordo M., Saggion H., Serra X. (In Press). Information Extraction for Knowledge Base Construction in the Music Domain. Journal on Knowledge & Data Engineering, Elsevier.

Relation Extraction (Tools)

ReVerb: OpenIE. Downloadable JAR. http://reverb.cs.washington.edu/

OpenIE: Successor of ReVerb. Downloadable JAR. http://openie.allenai.org/

DeepDive: Distant supervision. Installable python app.
http://deepdive.stanford.edu/

Outline

- Introduction to NLP
- Information Extraction
  - **Construction of Music Knowledge Bases**
  - Applications in MIR
- Topic Modeling
- Sentiment Analysis
- Lexical Semantics

Construction of Music KBS

Outline

- Motivation
- The Challenge of EL in the Music domain

  → ELMD and ELVIS

- Towards MKB Learning from Scratch

Motivation - Why you should care

- Structuring information in the Information Age is the big thing.

- Making sense of what people *say about music* has the potential to contribute dramatically to musicology and MIR.

    * Obtain knowledge automatically

    * Ask complex questions

    * Information Visualization

    * Improve navigation and personalization

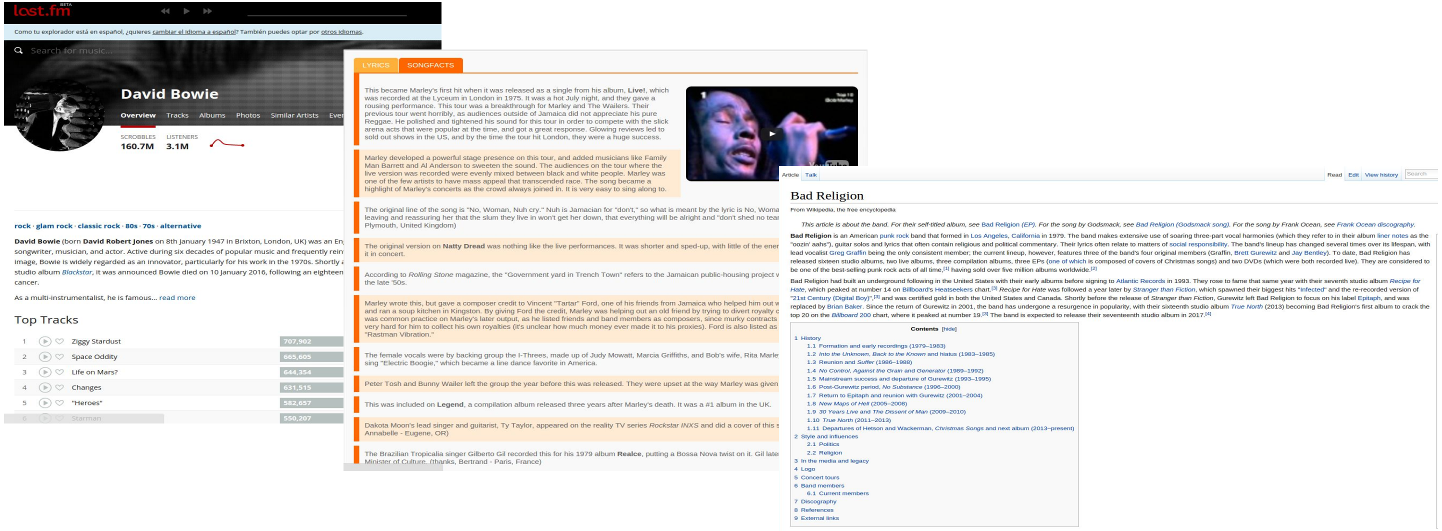
Motivation - Why you should care

- Structured information about music is incomplete
- Only popular artists and western music
- Only editorial and some biographical information

Motivation - Why you should care

- Huge amount of music information remains implicit in unstructured texts

    * Artists biographies, articles, reviews, web pages, user posts.



Challenges - Entity Linking

· **Entity Recognition.**

· Typical procedure: Gazetteers or knowledge repositories with musical information.

- Efficient in idiosyncratic and unambiguous cases: *The Symphony No. 9 in D minor.*
- But what if the name changes? For example, *The 9th is one of Beethoven's best.*
- One same mention may refer to different musical entities. E.g. *Carmen* the opera, and *Carmen* the opera's main character.
- Variability in musical entities. E.g. *The Rolling Stones* or *Their Satanic Majesties.*
- Musical entities with common names.
    - E.g. *Madonna* (artist or representation of Mary).

Challenges - Entity Linking

- Album and especially artist names get shortened in casual language.
- Album and artist names being the same.
- Generic software for Entity Linking don’t do well. Lack of sensitivity to musical text. Also, most of them exploit context, but this can be counterproductive.

Challenges - Entity Linking

System	Song	Album	Artist
Babelfy	**Carey**	**Debut**	John_Lennon
	Stephen	**Song_For**	Eminem
	Rap_Song	**Song_Of**	Paul_McCartney
Tagme	**The_Word**	Up	John_Lennon
	The_End	**When_We_On**	Do
	If	**Together**	Neil_Young
DBpedia Spotlight	Sexy_Sadie	The_Wall	Madonna
	Helter_Skelter	Let_It_Be	Eminem
	Cleveland_Rocks	Born_This_Way	Rihanna

ELMD: Entity Linking in the Music Domain

Oramas, S., Espinosa-Anke, L., Sordo, M., Saggion, H., & Serra, X. (2016). **ELMD: An Automatically Generated Entity Linking Gold Standard Dataset in the Music Domain**. In *In Proceedings of the 10th International Conference on Language Resources and Evaluation, LREC*.

ELMD: Entity Linking in the Music Domain

· We envisioned a text corpus annotated with a vast number of music entities (Album, Song, Artist and Record Label).

· While not all occurrences in text would be annotated, those who were should have very high Precision. Good for propagation, semi supervised learning, etc.

· We took advantage of artist biographies in last·fm

· And annotated dozens of thousands of entities with very high precision thanks to ELVIS!

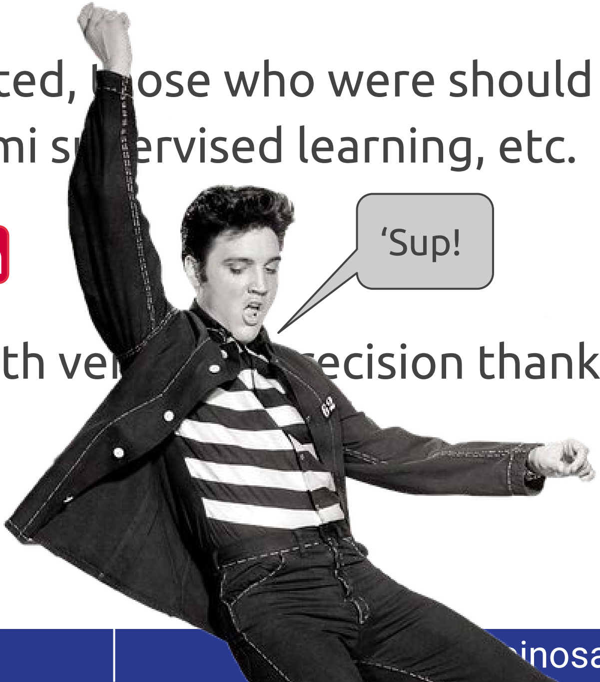
ELMD: Entity Linking in the Music Domain

- We envisioned a text corpus annotated with a vast number of music entities (Album, Song, Artist and Record Label).

- While not all occurrences in text would be annotated, t ose who were should have very high Precision. Good for propagation, semi s ervised learning, etc.

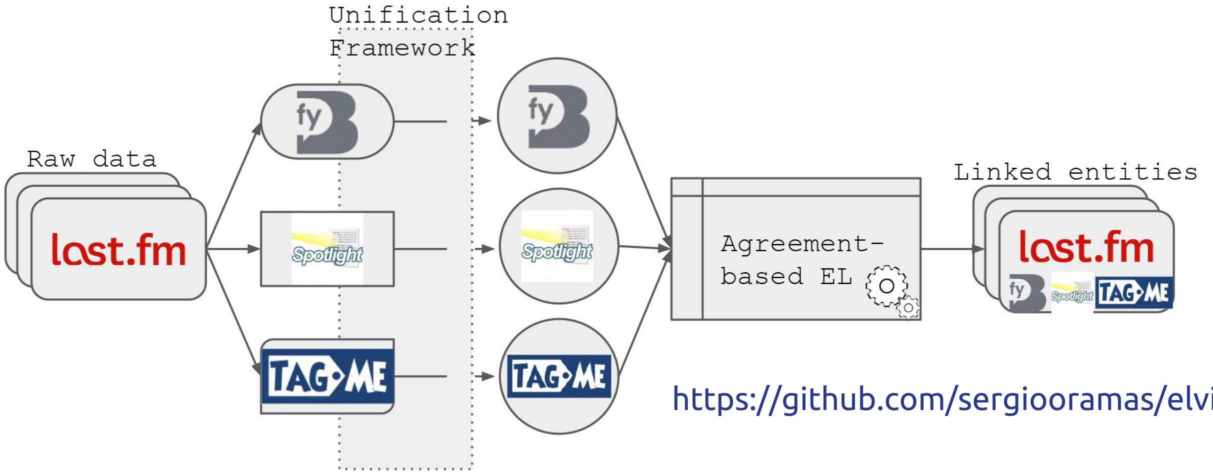
- We took advantage of artist biographies in last·fm

- And annotated dozens of thousands of entities with ve ecision thanks to ELVIS!



ELVIS: Entity Linking Voting and Integration System

· Assume agreement among generic tools can be leveraged to detect entities with *high precision*.



https://github.com/sergiooramas/elvis

ELMD: Entity Linking in the Music Domain

last.fm **Dataset**

* 13k artist biographies
* Collaborative effort
* Biographies are connected via 92,930 inner hyperlinks

- **ELMD: *Entity Linking in the Music Domain***
  * From hyperlinks to annotated named entities
  * Entities are then linked to DBpedia using ELVIS with 97% of precision

ELMD 2.0: Bigger and Better

- Novel entity disambiguation mapping to MusicBrainz.
- Existing annotations are heuristically propagated.
- Different output formats: JSON, XML GATE, NIF.
- 144,593 Annotations and 63,902 Entities.
- Full details and download available at: http://mtg.upf.edu/download/datasets/elmd

Towards MKB Learning from Scratch

Oramas, S., Espinosa-Anke, L., Sordo, M., Saggion, H., & Serra, X. (2016). **Information extraction for knowledge base construction in the music domain.** Data and Knowledge Engineering. *To appear.*

Towards MKB Learning from Scratch

· Starting from *songfacts.com* as a source for raw musical text, and after performing entity linking…

· The task lies now on how to leverage this information as the cornerstone of a music knowledge graph, the *backbone* of an MKB.

· The approach: Combine linguistically motivated rules over syntactic dependencies along with statistical evidence.

Towards MKB Learning from Scratch

· Shortest path doesn't always work

→ **Nile Rodgers** *told* NME that the first album he bought was 300 Impressions by **John Coltrane**.

⇒ **nile_rodgers** told that was impressions by **john_coltrane**

· Consider special cases of:

* Reported speech ("say", "tell", "express")

* Enforce certain syntactic relations between entity and first relation word.

* etc

Towards MKB Learning from Scratch

- **Relation Clustering: Syntactic Dependencies + Type Filtering**

Cluster Pattern	Typed cluster pattern	Relation triple
was written by	song was written by artist	**song** was written by artist **artist**
		song was written by composer **artist**
		song was written by **artist**
	album was written by artist	**album** was written by frontman **artist**
		album was written by guitarist **artist**
		album was written by artist **artist**
		album was written by newcomer **artist**

Towards MKB Learning from Scratch

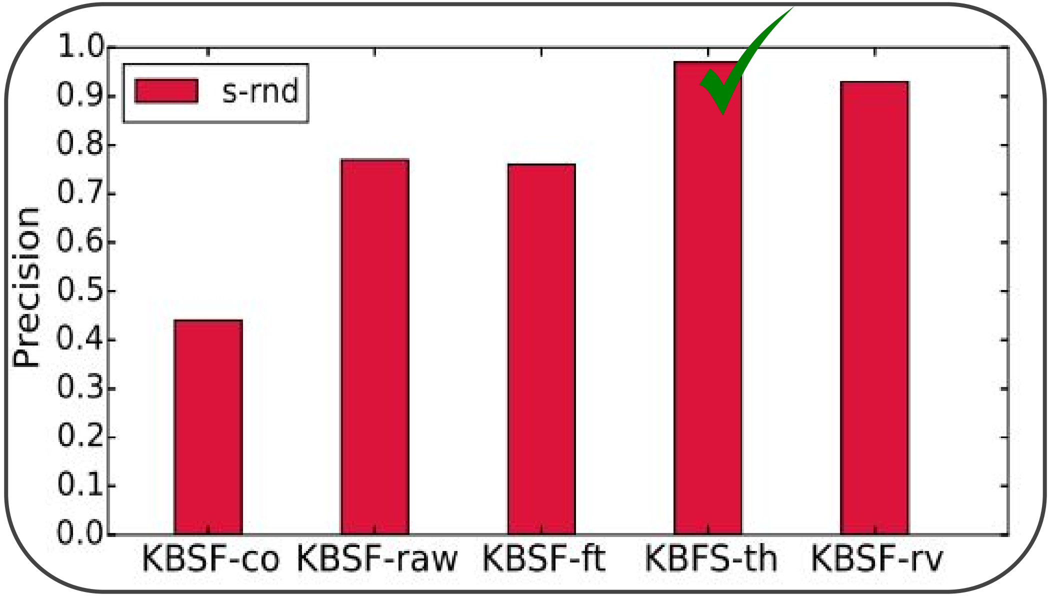
- **Relation Scoring**

- The relevance of a cluster may be inferred by the number and proportion of triples it encodes, and whether these are evenly distributed.

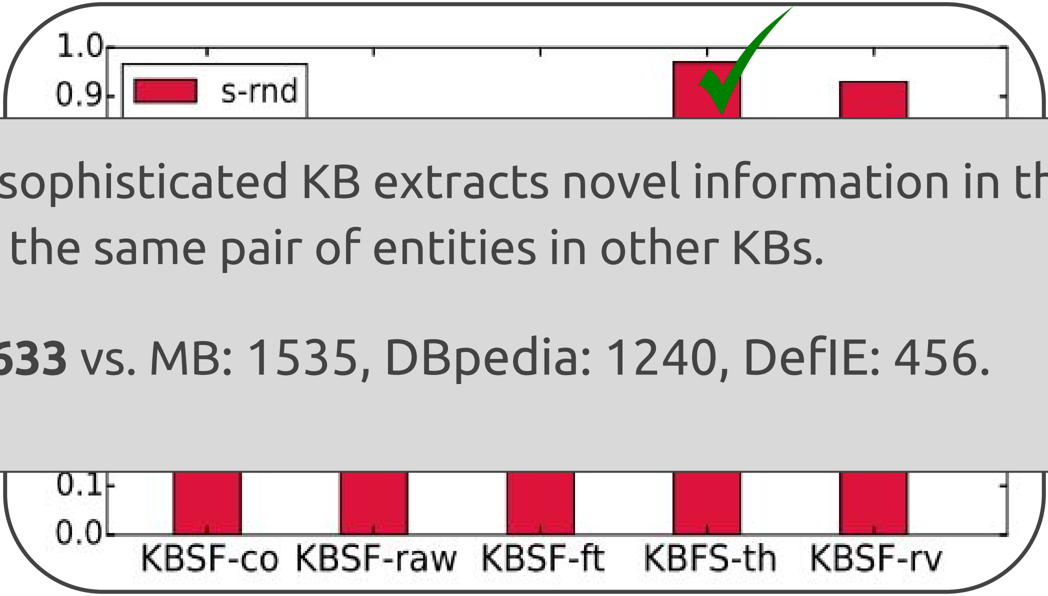
- Degree of specificity. ⇒ <$artist_d$, *performed_with*, $artist_r$>

- Frequency, lenght and fluency. Reward those relations which preserve the original sentence' word order.

Towards MKB Learning from Scratch



Towards MKB Learning from Scratch



Our most sophisticated KB extracts novel information in the form of triples for the same pair of entities in other KBs.

Our KB: **3633** vs. MB: 1535, DBpedia: 1240, DefIE: 456.

Towards MKB Learning from Scratch

Towards MKB Learning from Scratch

- **Bruce Springsteen** *covered* **Jersey Girl**

Towards MKB Learning from Scratch

- **Bruce Springsteen** *covered* **Jersey Girl**
- **Bruce Springsteen** *player* **Clarence Clemons**

Towards MKB Learning from Scratch


- **Bruce Springsteen** *covered* **Jersey Girl**
- **Bruce Springsteen** *player* **Clarence Clemons**
- **Hair (Lady Gaga)** *features* **Clarence Clemons**

Towards MKB Learning from Scratch

- Bruce Springsteen / Jersey Girl

- Bruce Sp lemons
- Hair (Lady Clemons

Towards MKB Learning from Scratch

- **Conclusion**

- Lots of unstructured information about music in the form of natural language

- We have barely scratched the surface. No Social Networks, no Wikipedia, no lyrics, no subtitles…

- Potential for improving MIR and musicological resources by integrating automatically acquired knowledge via Natural Language Processing.

References

Yago: Suchanek, F. M., Kasneci, G., & Weikum, G. (2007, May). Yago: a core of semantic knowledge. In Proceedings of the 16th international conference on World Wide Web (pp. 697-706). ACM.

Wikidata: Vrandečić, D., & Krötzsch, M. (2014). Wikidata: a free collaborative knowledgebase. Communications of the ACM, 57(10), 78-85.

Freebase: Bollacker, K., Evans, C., Paritosh, P., Sturge, T., & Taylor, J. (2008, June). Freebase: a collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data (pp. 1247-1250). ACM.

MusicBrainz: Swartz, A. (2002). Musicbrainz: A semantic web service. Intelligent Systems, IEEE, 17(1), 76-77.

DBpedia: Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007). Dbpedia: A nucleus for a web of open data (pp. 722-735). Springer Berlin Heidelberg.

DBpedia Spotlight: Mendes, P. N., Jakob, M., García-Silva, A., & Bizer, C. (2011, September). DBpedia spotlight: shedding light on the web of documents. In Proceedings of the 7th International Conference on Semantic Systems (pp. 1-8). ACM.

BabelNet: Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. Artificial Intelligence, 193, 217-250.

Babelfy: Moro, A., Raganato, A., & Navigli, R. (2014). Entity linking meets word sense disambiguation: a unified approach. Transactions of the Association for Computational Linguistics, 2, 231-244.

Tagme: Ferragina, P., & Scaiella, U. (2010, October). Tagme: on-the-fly annotation of short text fragments (by wikipedia entities). In Proceedings of the 19th ACM international conference on Information and knowledge management (pp. 1625-1628). ACM.

Motivation - Why you should care

- But what are Knowledge Bases?
- Structured representations of knowledge stored, usually, in the form of graphs.
- They can be created **manually** (WordNet).
- Or automatically from **semi-structured knowledge** (DBpedia)
- Or automatically by **unifying knowledge** into one resource (BabelNet)
- Or automatically **reading the web** (NELL)

Outline

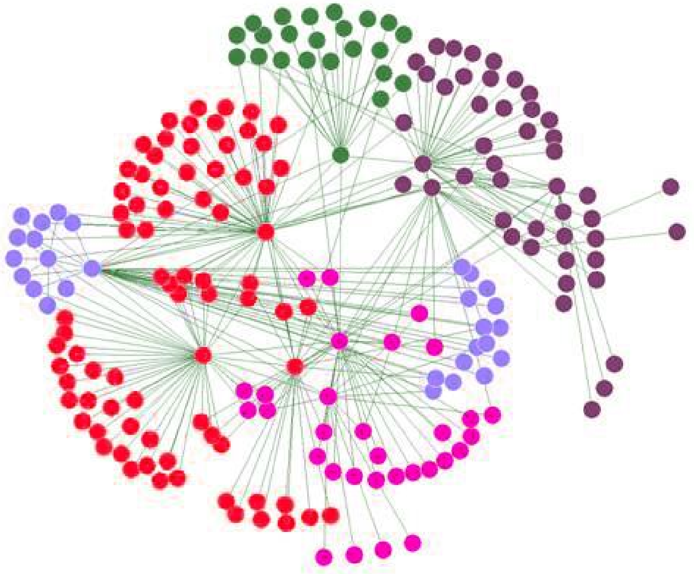
- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - **Applications in MIR**
- Topic Modeling
- Sentiment Analysis
- Lexical Semantics

Use semantic information extracted from text

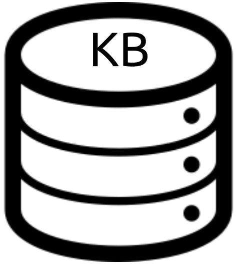
Approach: Create a knowledge representation with graph structure and then apply graph-based methodologies.

Several **types of graphs**:

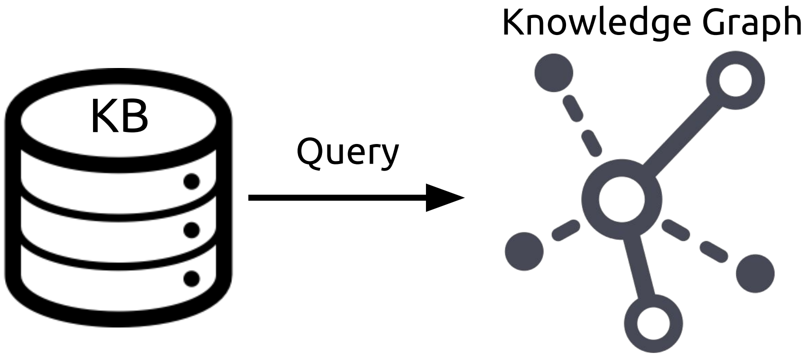
- Knowledge Graph
- Graph of Entities
- Semantically Enriched Graph



Graphs



Graphs



Knowledge Graph

Wilco
dbo:bandMember -> dbr:Jeff_Tweedy
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:Illinois

Son Volt
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:St._Louis,_Missouri
dbo:recordLabel -> dbr:Warner_Bros._Records

Knowledge Graph

Wilco
dbo:bandMember -> dbr:Jeff_Tweedy
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:Illinois

Son Volt
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:St._Louis,_Missouri
dbo:recordLabel -> dbr:Warner_Bros._Records

DBpedia

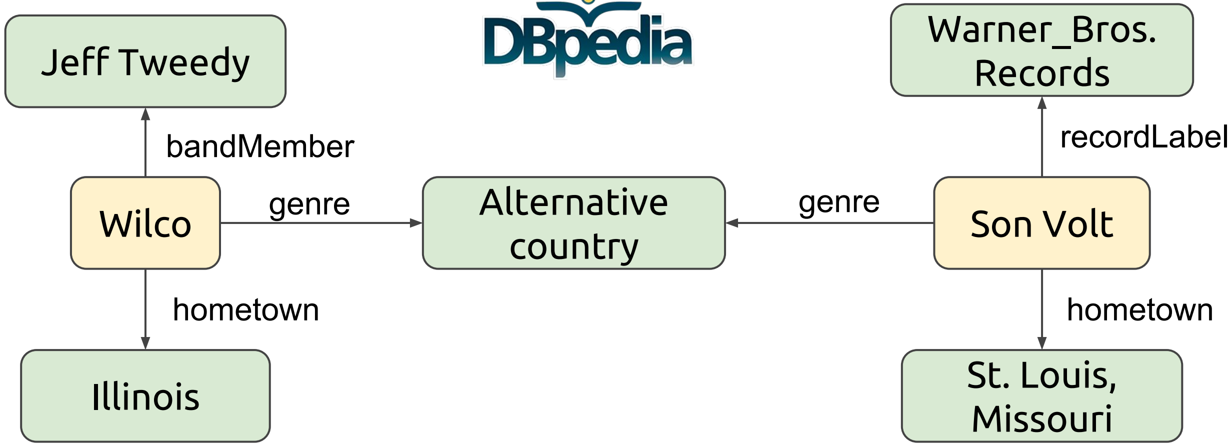
Wilco

Son Volt

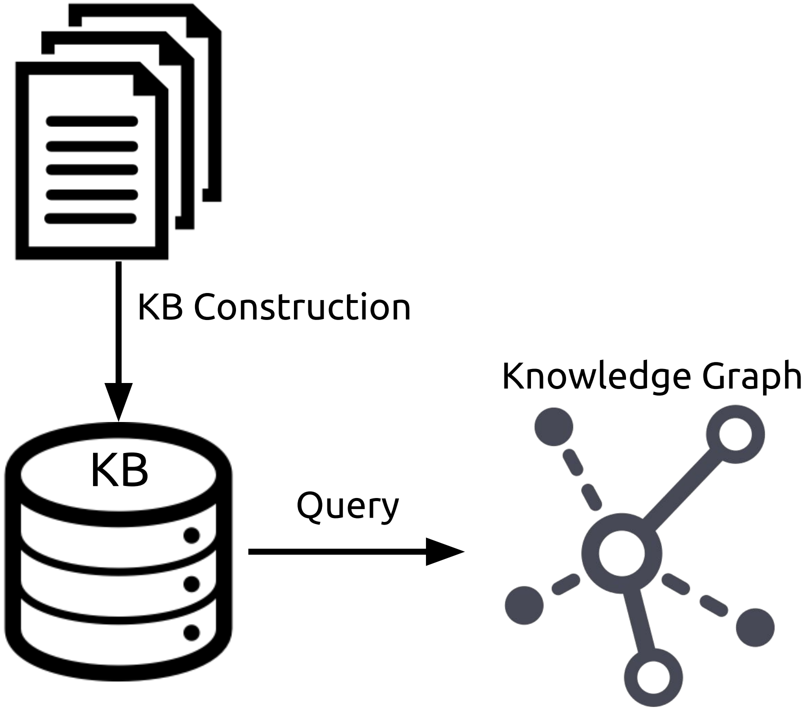
Knowledge Graph

Wilco
dbo:bandMember -> dbr:Jeff_Tweedy
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:Illinois

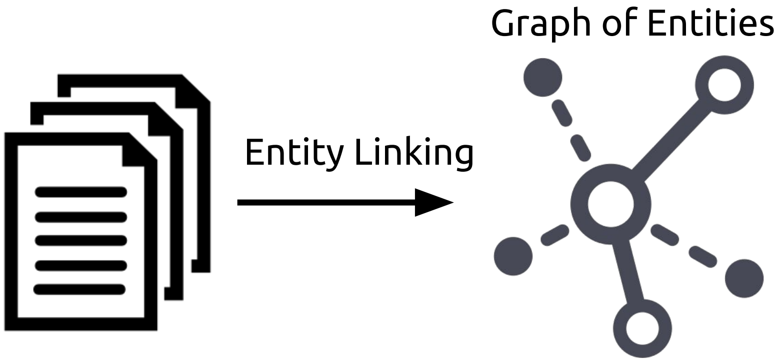
Son Volt
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:St._Louis,_Missouri
dbo:recordLabel -> dbr:Warner_Bros._Records



Graphs



Graphs



Graph of Entities

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

Graph of Entities

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

Entity Linking

Graph of Entities

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

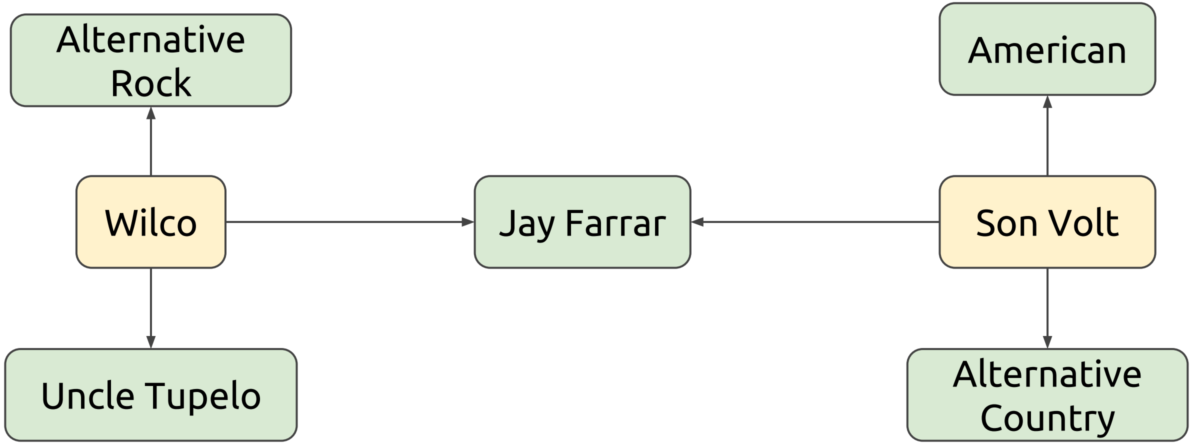
Wilco

Son Volt

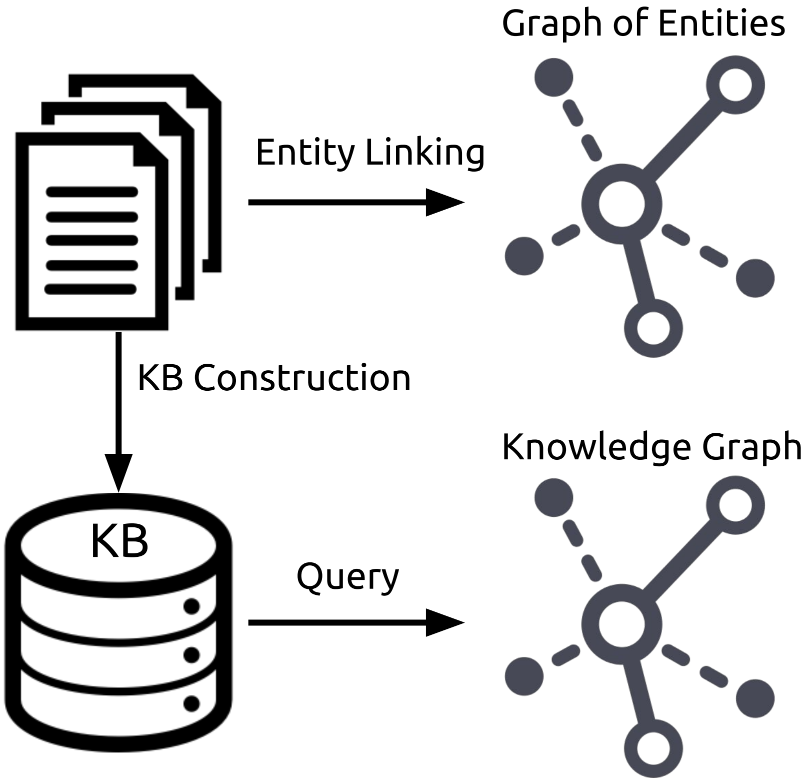
Graph of Entities

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

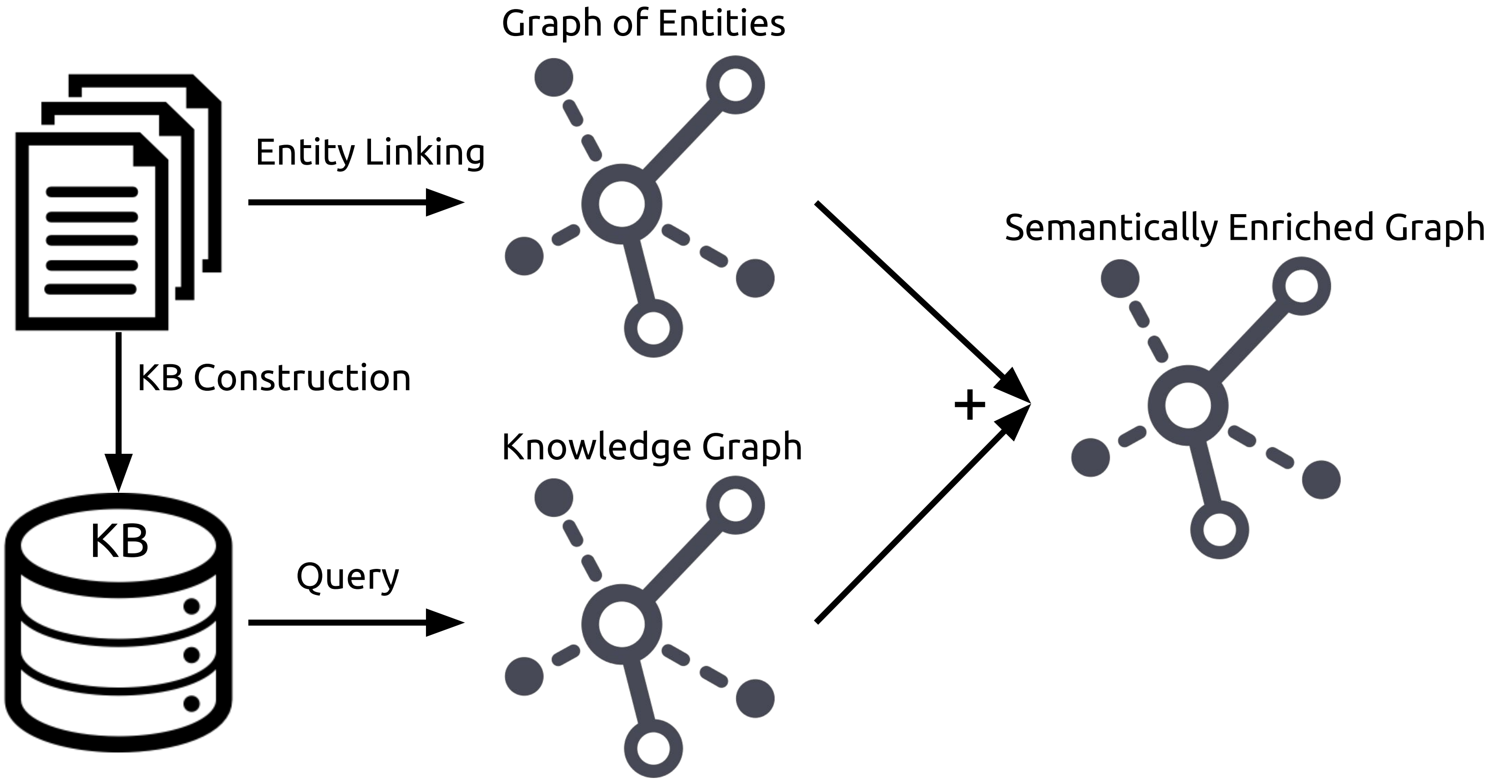


Graphs



Graph of Entities
Entity Linking
KB Construction
Knowledge Graph
KB
Query

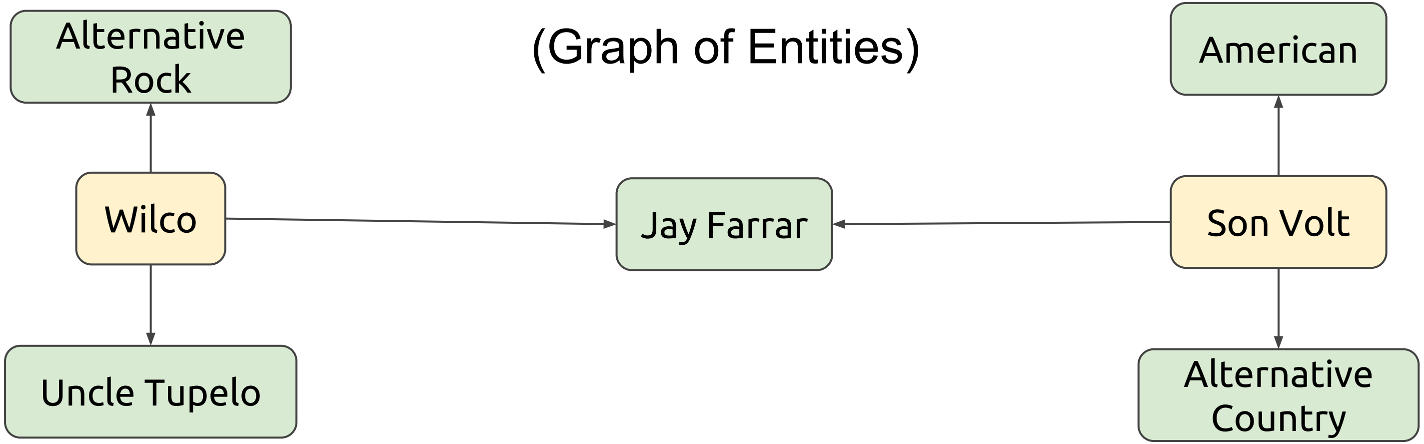
Graphs


Graph of Entities
Entity Linking
Semantically Enriched Graph
KB Construction
+
Knowledge Graph
KB
Query

Semantically Enriched Graph

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.



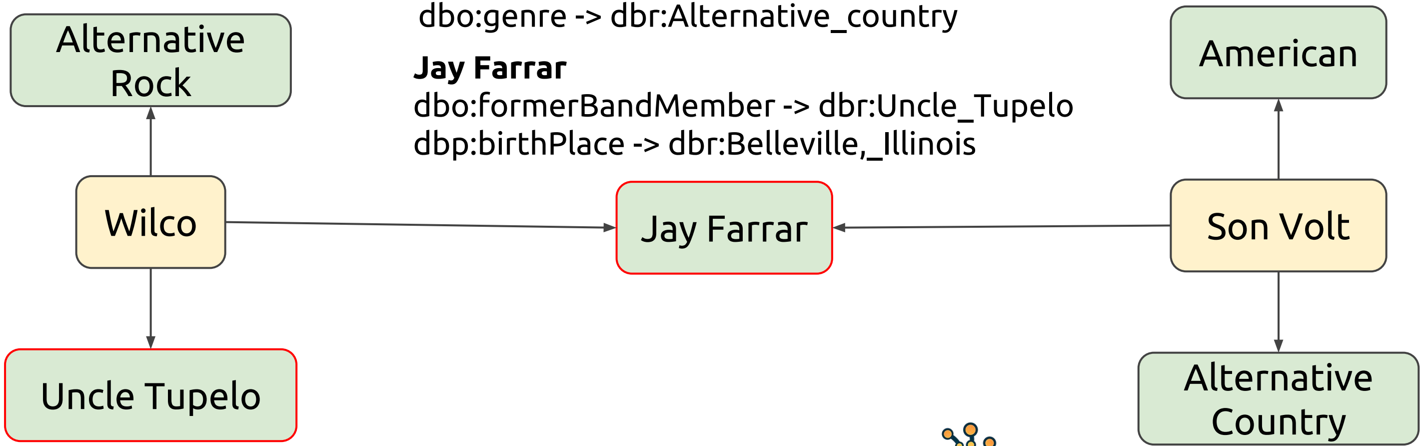
Semantically Enriched Graph

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

Uncle Tupelo
dbo:hometown -> dbr:Belleville,_Illinois
dbo:genre -> dbr:Alternative_country

Jay Farrar
dbo:formerBandMember -> dbr:Uncle_Tupelo
dbp:birthPlace -> dbr:Belleville,_Illinois



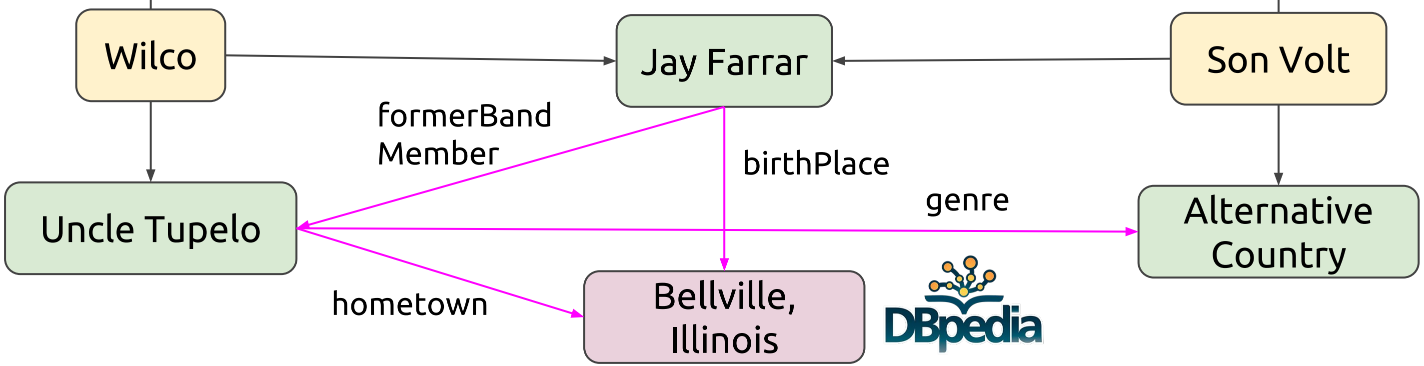
Semantically Enriched Graph

Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

Uncle Tupelo
dbo:hometown -> dbr:Belleville,_Illinois
dbo:genre -> dbr:Alternative_country

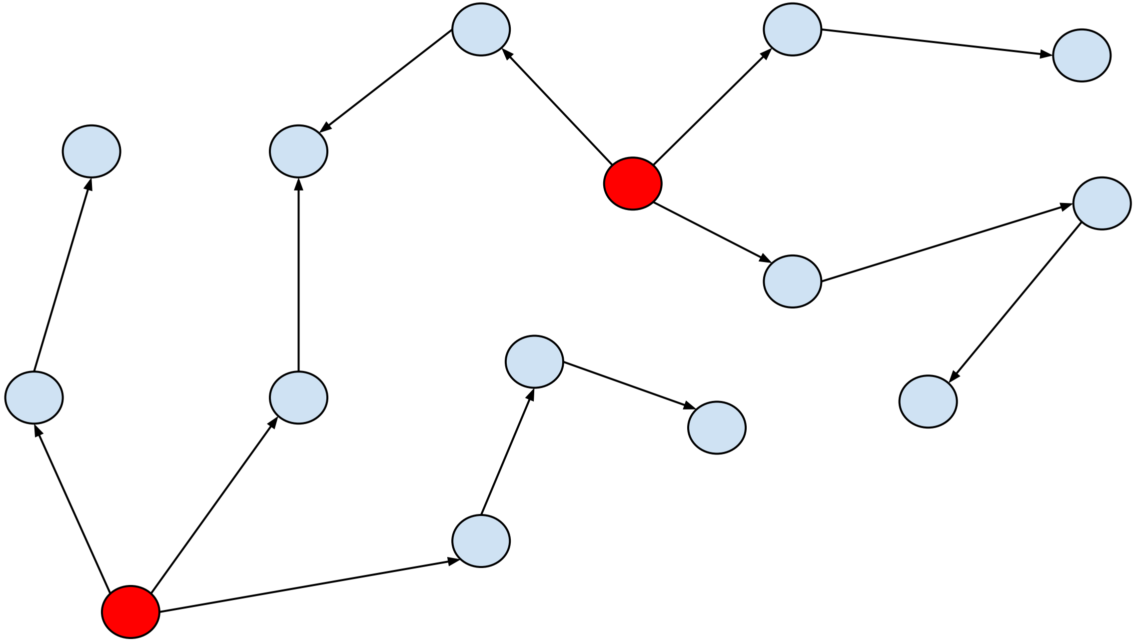
Jay Farrar
dbo:formerBandMember -> dbr:Uncle_Tupelo
dbp:birthPlace -> dbr:Belleville,_Illinois



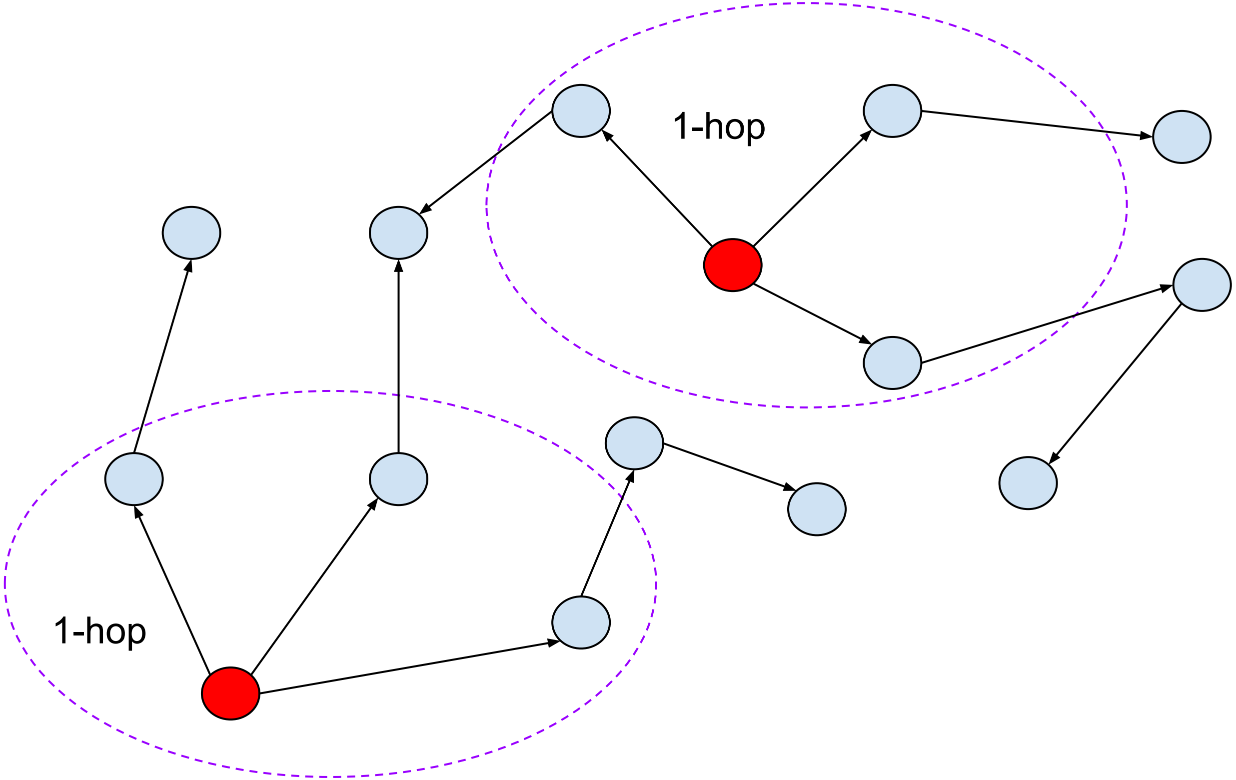
Exploiting Graph Information

- Measure similarity between entities in a graph
- Embed graphs into linear vectors
- Use of Knowledge Graphs in Music Recommendation
- Computing relevance of the entities in a graph
- Analytics
- Visualization of Graphs

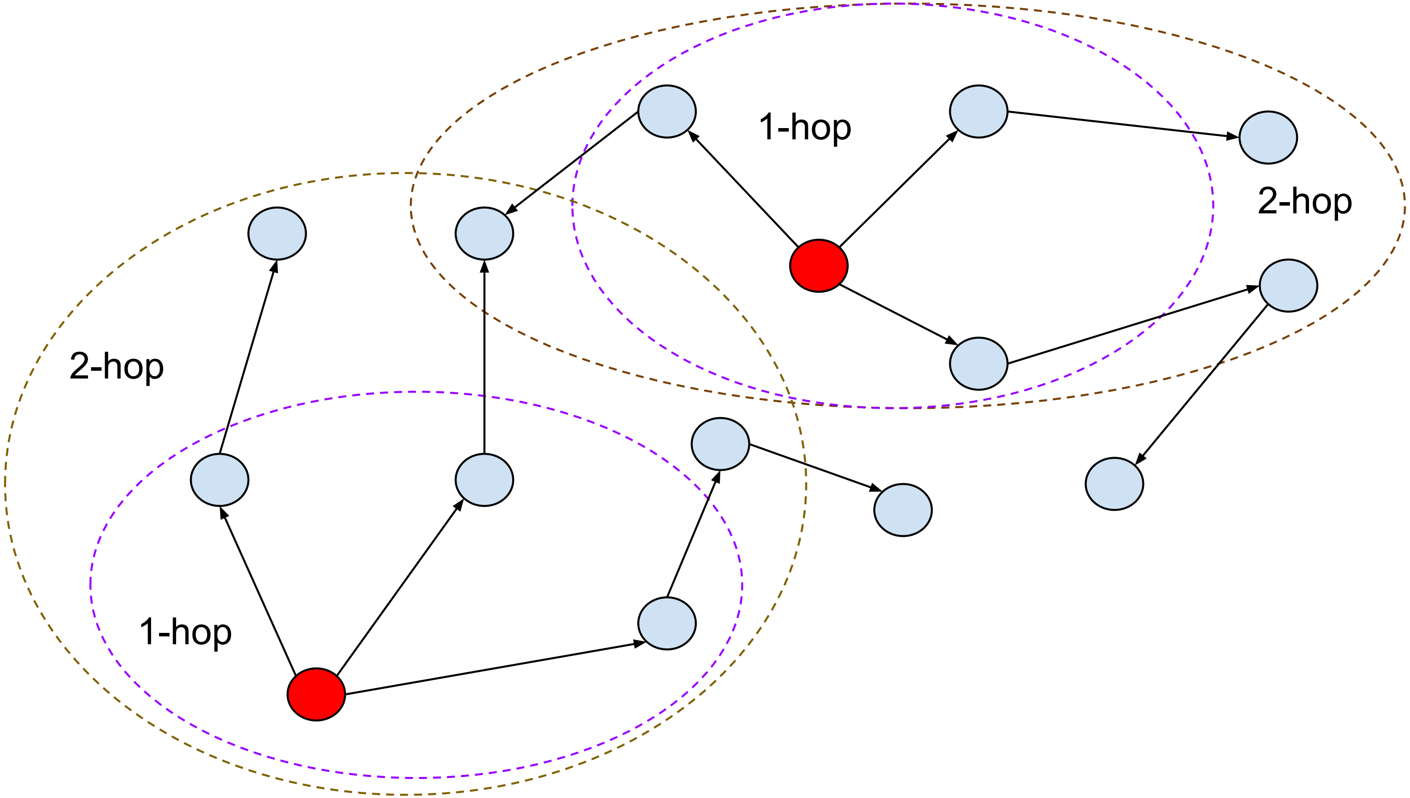
h-hop Item Neighborhood Graph



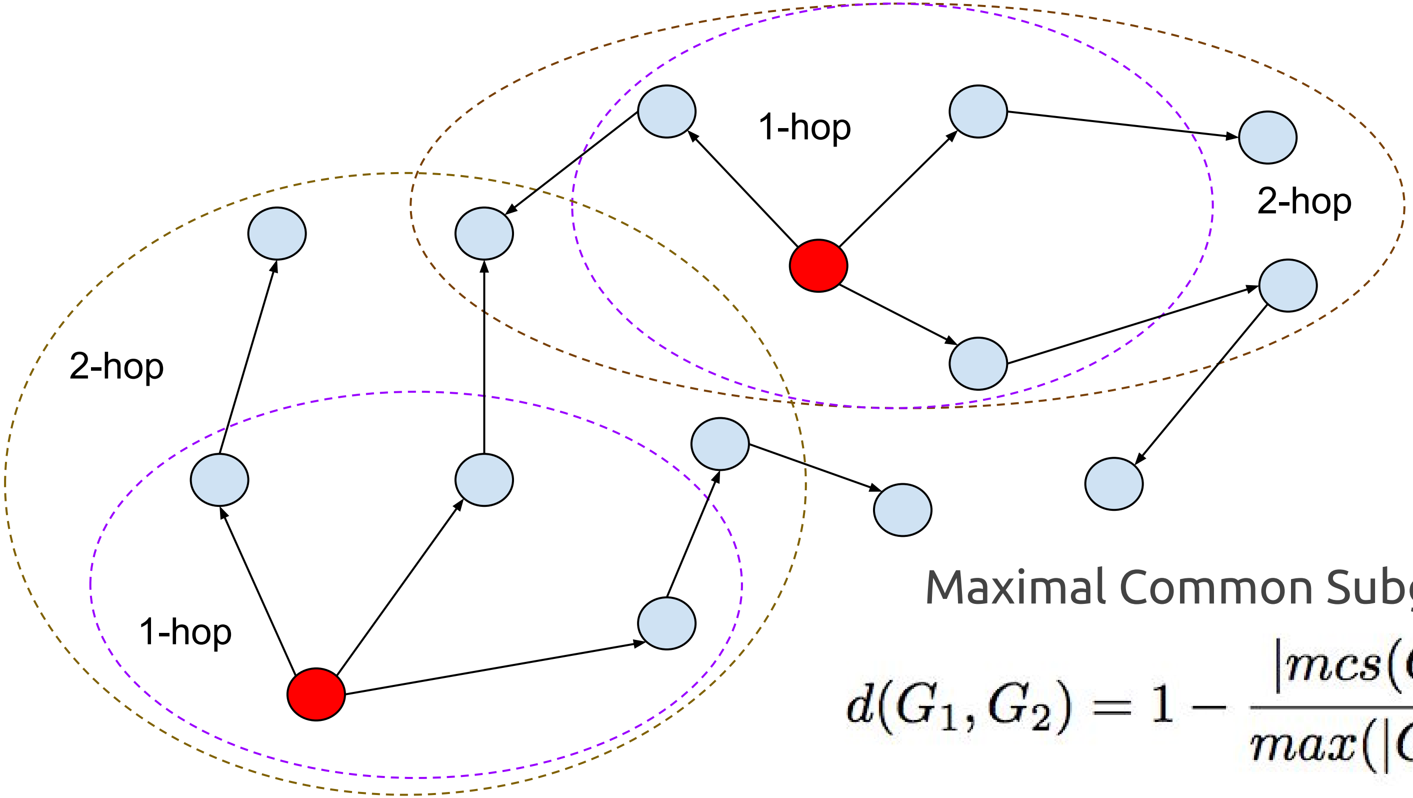
h-hop Item Neighborhood Graph



h-hop Item Neighborhood Graph



Artist Similarity



Maximal Common Subgraph

$$d(G_1, G_2) = 1 - \frac{|mcs(G_1, G_2)|}{max(|G_1|, |G_2|)}$$

Artist Similarity

Oramas S., Sordo M., Espinosa-Anke L., & Serra X. (2015). *A Semantic-based approach for Artist Similarity*. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

- Artist biographies gathered from Last.fm
- Entity Linking tool used: Babelfy
- Build different knowledge graphs
- Two Experiments:
  - MIREX: 188 artists, MIREX Audio and Music Similarity evaluation dataset
  - Last.fm API: 2,336 artists, Last.fm API similarity

Evaluation dataset: http://mtg.upf.edu/download/datasets/semantic-similarity

Artist Similarity

Evaluation

Approach variants	Precision@N		nDCG@N	
	N=5	N=10	N=5	N=10
LSA	0.090	0.088	0.233	0.269
RG MCS 1-hop	0.055	0.083	0.126	0.149
AE MCS	0.124	0.200	0.184	0.216
AE-FT MCS	0.136	0.201	0.224	0.260
AEC MCS 1-hop	0.152	0.224	0.277	0.297
AEC-FT MCS 1-hop	**0.160**	**0.242**	**0.288**	**0.317**

Text based approach (BoW)

Knowledge Graph from Extracted KB

Graph of Entities

Semantically Enriched Graph

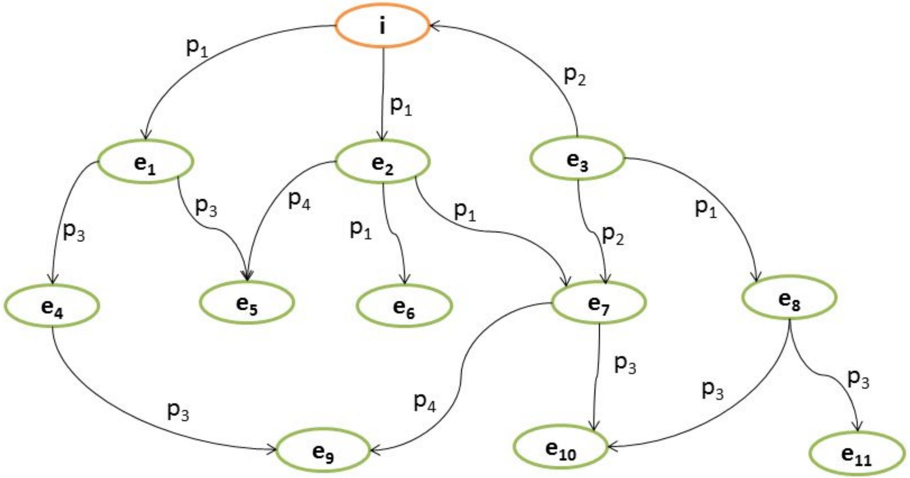
Graph Embeddings

Encode graph information into a vector space

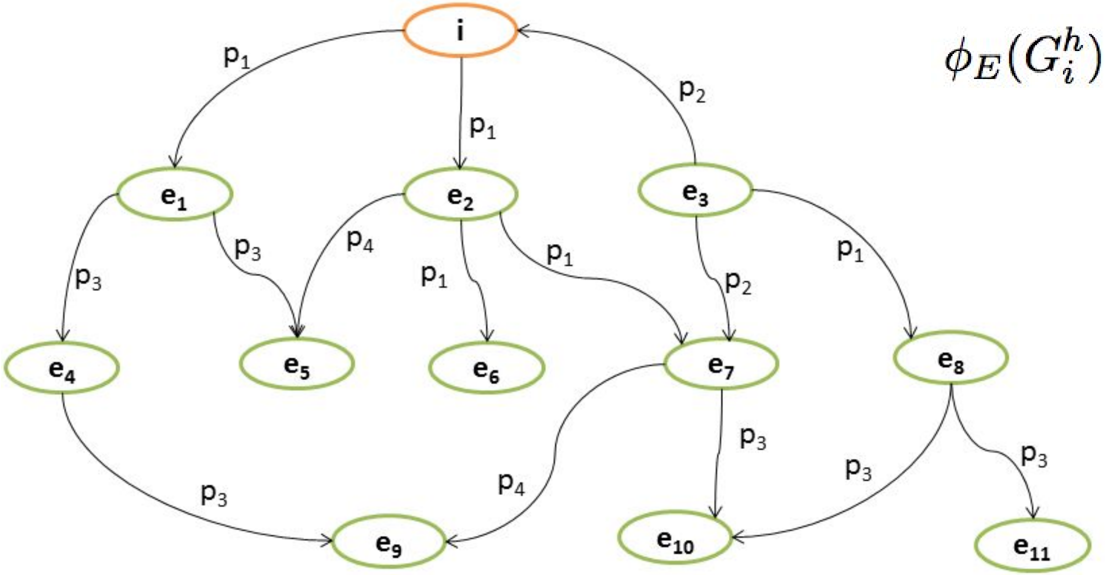
2 different **embedding approaches**

- Entity-based
- Path-based

Useful for recommendation



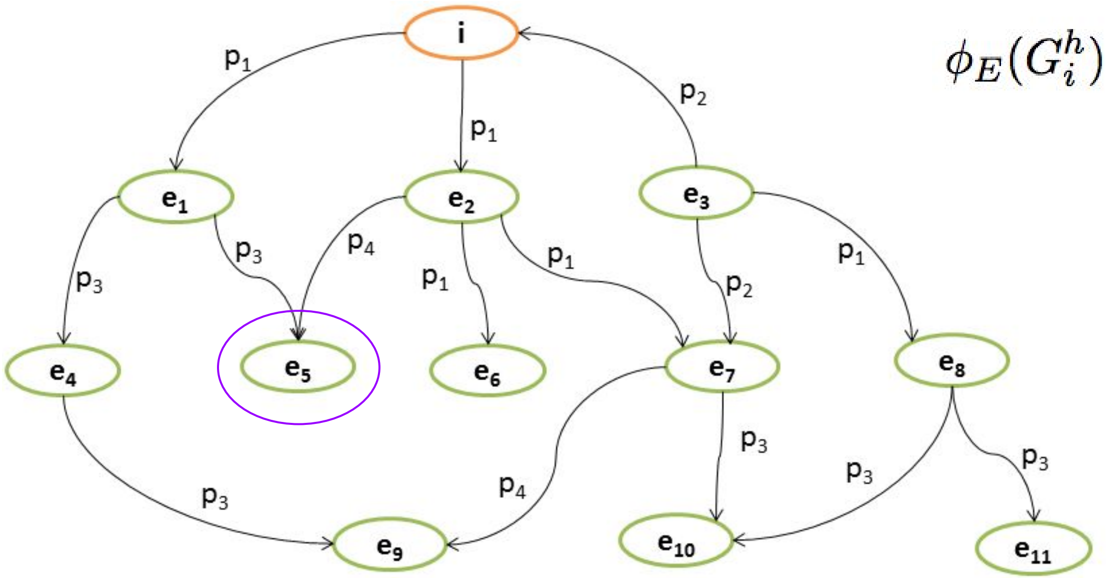
Entity-based Item Neighborhood Mapping



$$\phi_E(G_i^h) = (w_{i,e_1}, w_{i,e_2}, ...w_{i,e_m}, ..., w_{i,e_t})$$

- One feature per entity
- Weight according to:
  - Distance to item
  - Number of in-links

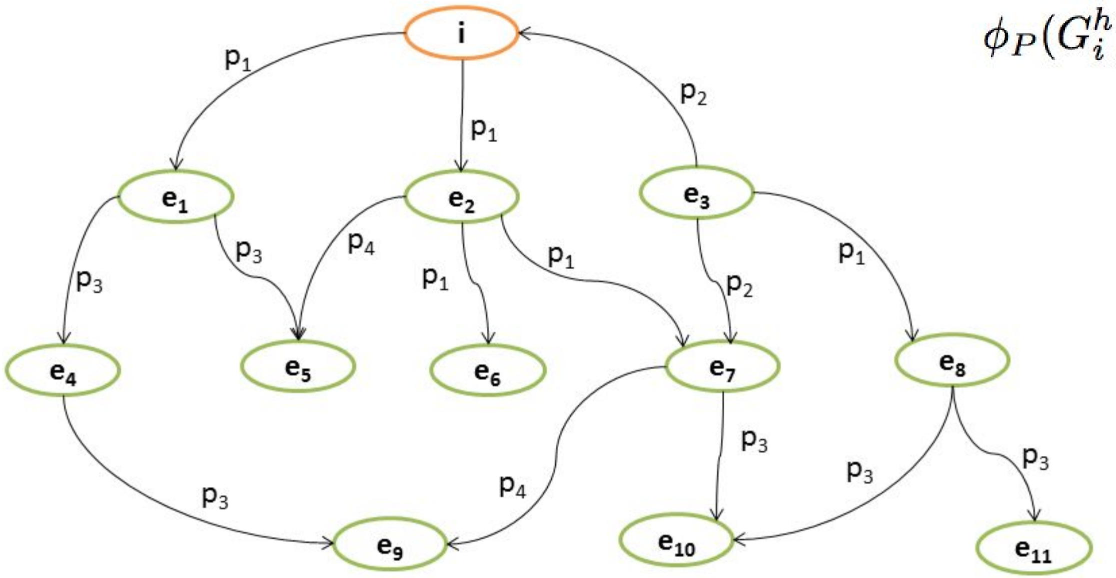
Entity-based Item Neighborhood Mapping



$$\phi_E(G_i^h) = (w_{i,e_1}, w_{i,e_2}, ...w_{i,e_m}, ..., w_{i,e_t})$$

- One feature per entity
- Weight according to:
  - Distance to item
  - Number of in-links

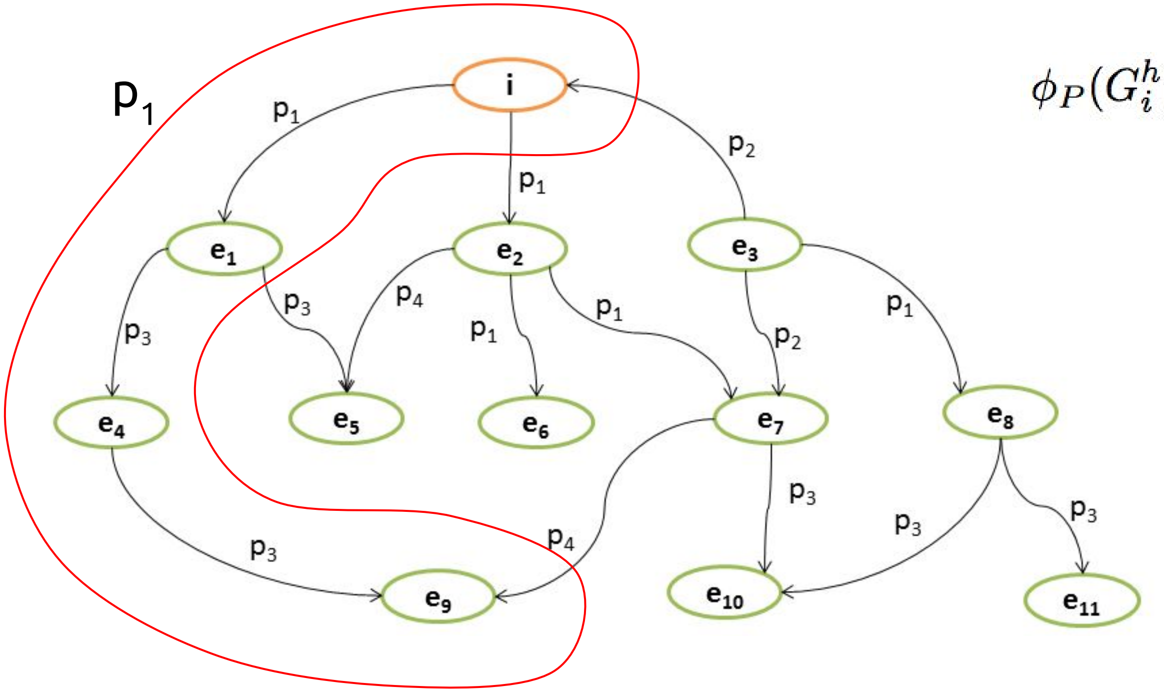
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

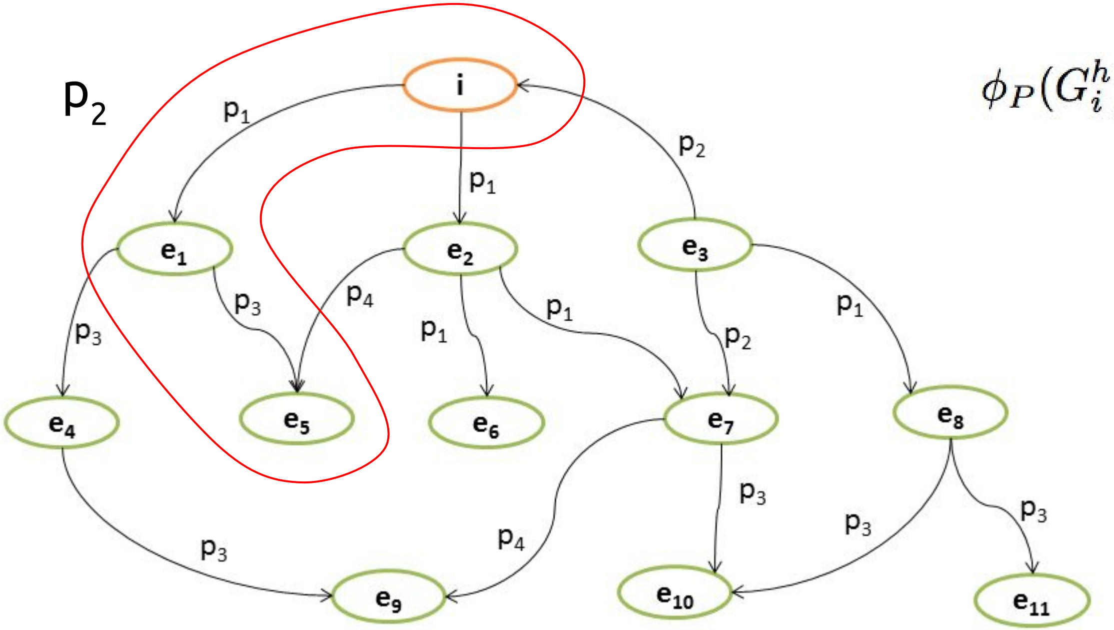
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, ...w_{i,p_m^*}, ..., w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

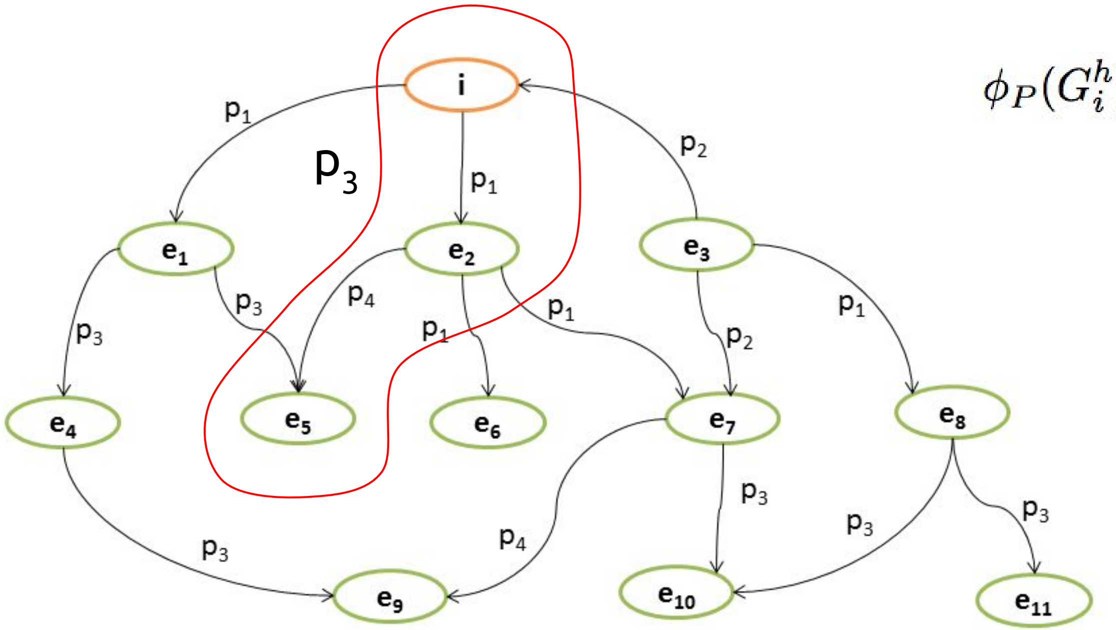
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, ...w_{i,p_m^*}, ..., w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

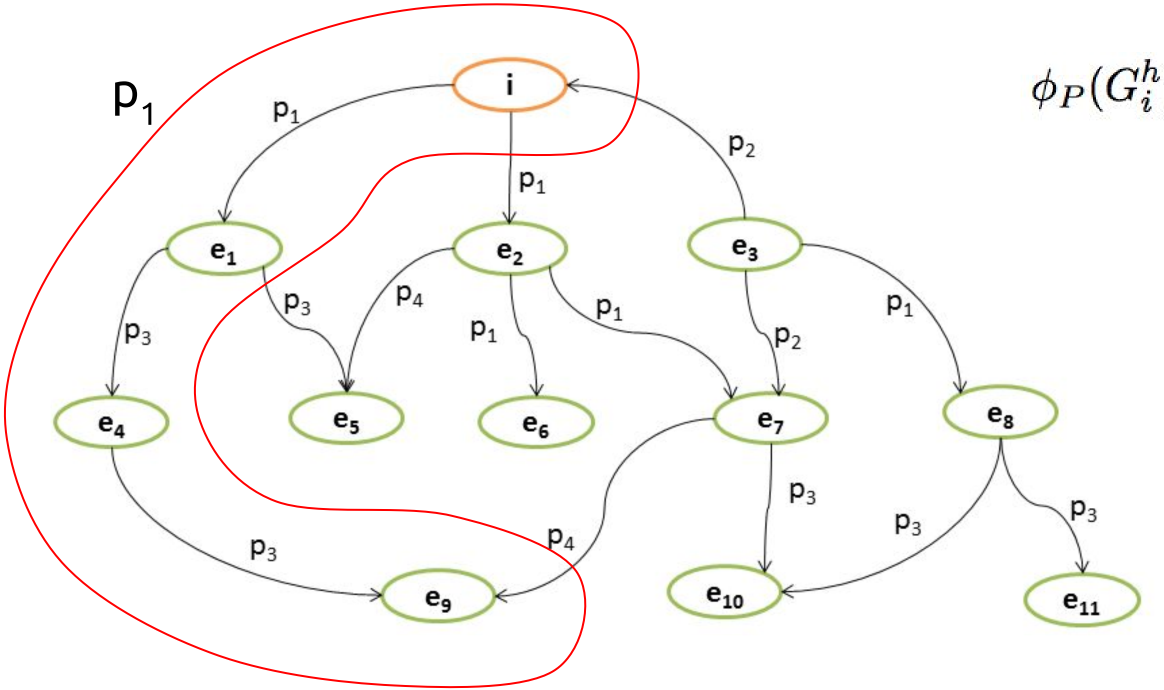
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

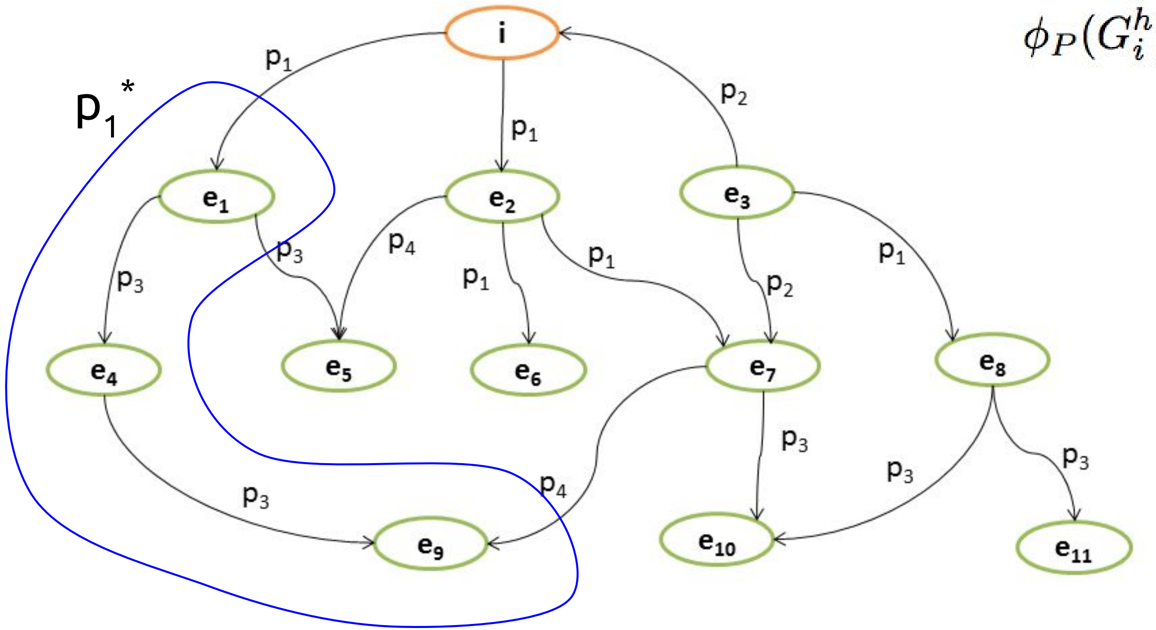
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

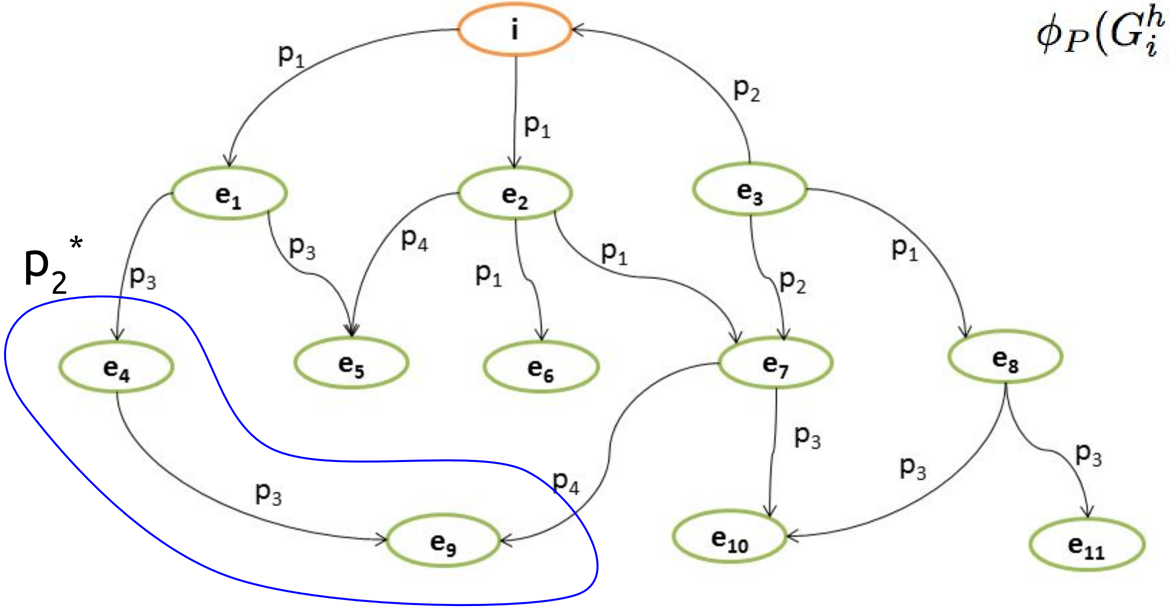
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

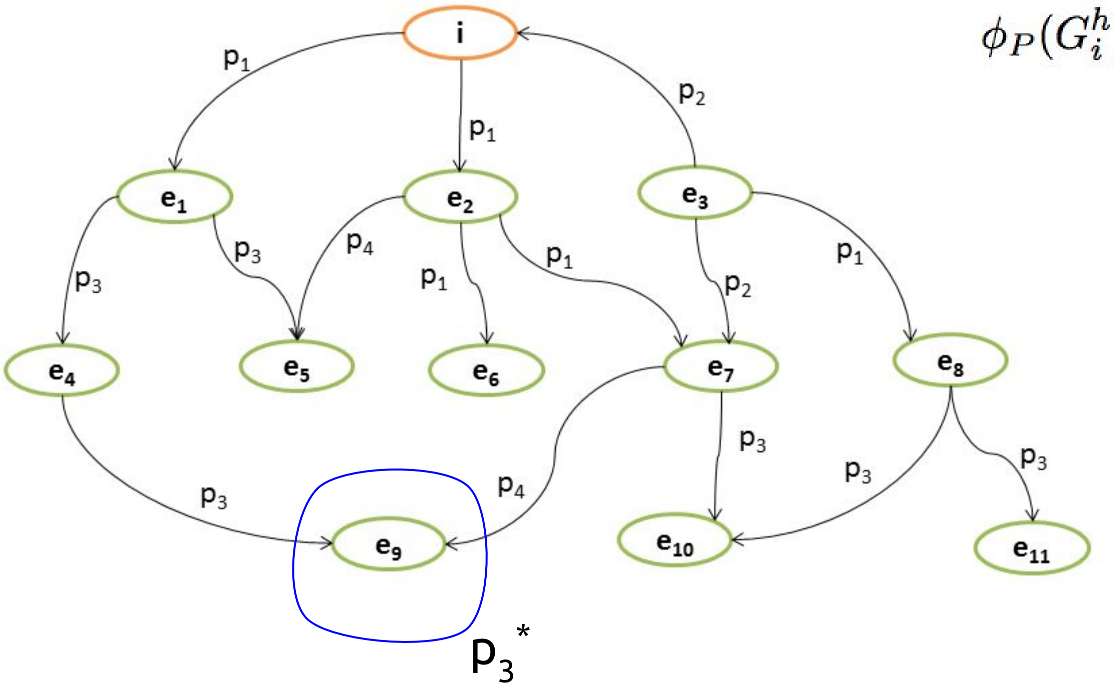
Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Item Neighborhood Mapping



$$\phi_P(G_i^h) = (w_{i,p_1^*}, w_{i,p_2^*}, \ldots w_{i,p_m^*}, \ldots, w_{i,p_t^*})$$

- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

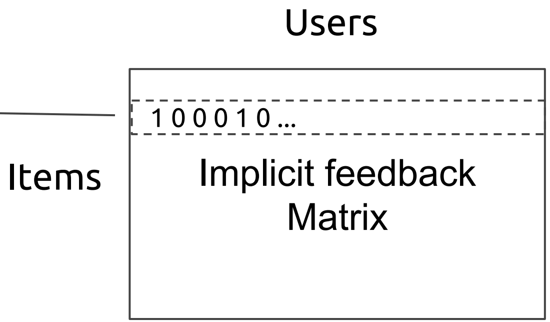
Music Recommendation

Hybrid approach: Knowledge Graph features + collaborative features

Collaborative features vector:

$$\phi_{col}(i) = (w_{i,u_1}, w_{i,u_2}, \ldots, w_{i,u_1})$$

where w_{i},u_{j} = 1 if user u_j interacted with item i



Music Recommendation

Aggregation of features

Item vector	Knowledge Graph vector	Collaborative vector

Train a regression model on every user

Oramas S., Ostuni V. C., Di Noia T., Serra, X., & Di Sciascio E. (2016). Music and Sound Recommendation with Knowledge Graphs. ACM Transactions on Intelligent Systems and Technology.

Source code: https://github.com/sisinflab/lodreclib

Music Recommendation

Two **experiments**:

- Sounds Recommendation
  - Freesound tags and descriptions + Implicit feedback (downloads)
  - 21,552 items and 20,000 users
- Music Recommendation
  - Last.fm tags and Songfacts descriptions + Implicit feedback (Last.fm listening habits)
  - 8,640 items and 5,199 users

Datasets: http://mtg.upf.edu/download/datasets/knowledge-graph-rec

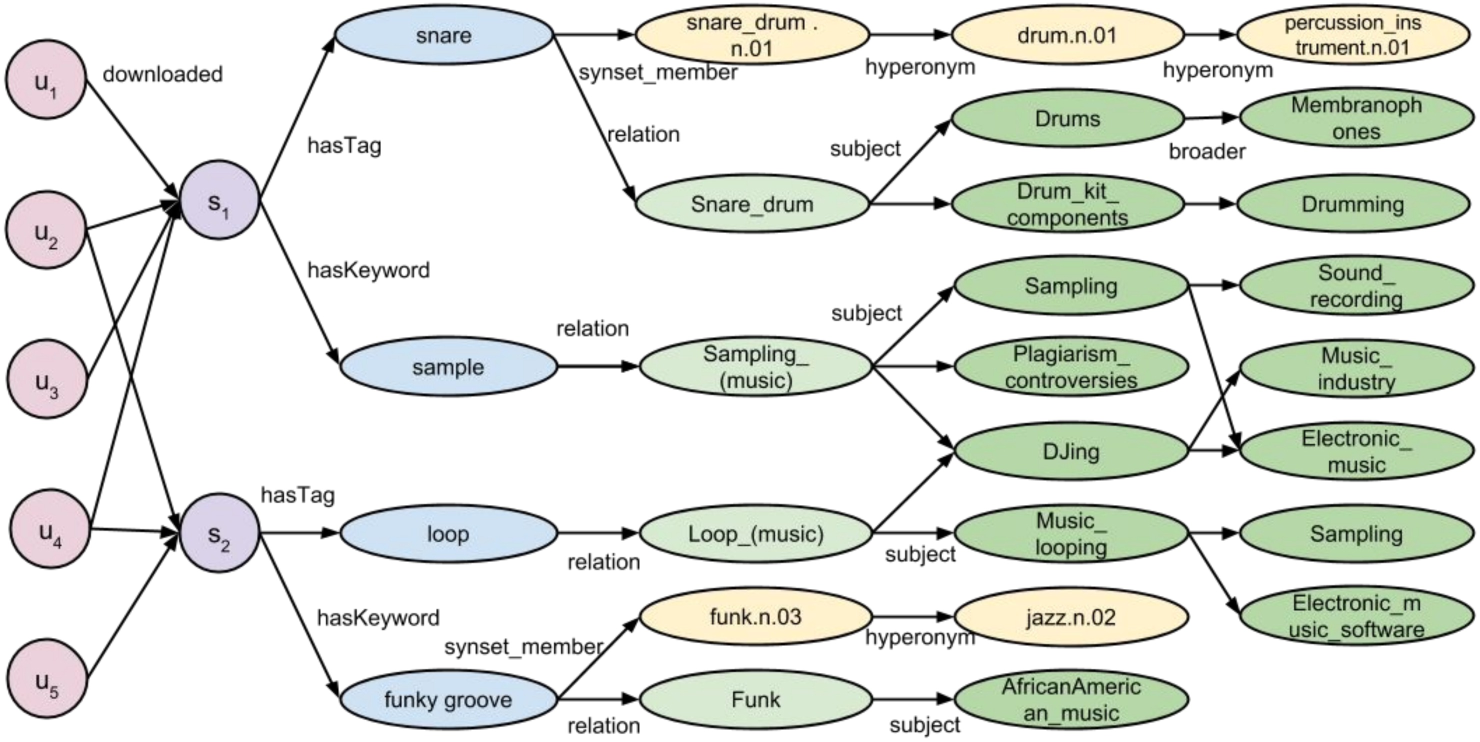
Music Recommendation

Knowledge Graph approach

- Semantically Enriched Graph over tags and text descriptions
- Using Babelfy for Entity Linking
- Using Wikipedia categories and WordNet synsets and hypernymy relations for semantic expansion

dataset	items	avg. tags	avg. keywords	resources	synsets	categories
Freesound	21,552	6.44	11.36	16,407	20,034	54,419
Last.fm	8,640	42.09	77.33	46,109	27,708	96,942

Music Recommendation



Music Recommendation

KG features	Collab features	P@10	R@10	EBN@10	ADiv@10
Entity-based	si	**0.118**	**0.067**	2.426	0.361
Path-based	si	0.111	0.061	1.618	0.532
Path-based	no	0.049	0.028	**0.369**	**0.670**
-	si	0.110	0.062	2.890	0.181
VSM	si	0.116	0.066	2.621	0.305
Audio Sim	no	0.004	0.002	0.382	0.044

EBN: Entropy-based Novelty
ADiv: Aggregated Diversity

Music Recommendation

KG features	Collab features	P@10	R@10	EBN@10	ADiv@10
Entity-based	si	**0.118**	**0.067**	2.426	0.361
Path-based	si	0.111	0.061	1.618	0.532
Path-based	no	0.049	0.028	**0.369**	**0.670**
-	si	0.110	0.062	2.890	0.181
VSM	si	0.116	0.066	2.621	0.305
Audio Sim	no	0.004	0.002	0.382	0.044

EBN: Entropy-based Novelty
ADiv: Aggregated Diversity

Music Recommendation

KG features	Collab features	P@10	R@10	EBN@10	ADiv@10
Entity-based	si	**0.118**	**0.067**	2.426	0.361
Path-based	si	0.111	0.061	1.618	0.532
Path-based	no	0.049	0.028	**0.369**	**0.670**
-	si	0.110	0.062	2.890	0.181
VSM	si	0.116	0.066	2.621	0.305
Audio Sim	no	0.004	0.002	0.382	0.044

EBN: Entropy-based Novelty
ADiv: Aggregated Diversity

Music Recommendation

KG features	Collab features	P@10	R@10	EBN@10	ADiv@10
Entity-based	si	**0.118**	**0.067**	2.426	0.361
Path-based	si	0.111	0.061	1.618	0.532
Path-based	no	0.049	0.028	**0.369**	**0.670**
-	si	0.110	0.062	2.890	0.181
VSM	si	0.116	0.066	2.621	0.305
Audio Sim	no	0.004	0.002	0.382	0.044

EBN: Entropy-based Novelty
ADiv: Aggregated Diversity

Music Recommendation

KG features	Collab features	P@10	R@10	EBN@10	ADiv@10
Entity-based	si	**0.118**	**0.067**	2.426	0.361
Path-based	si	0.111	0.061	1.618	0.532
Path-based	no	0.049	0.028	**0.369**	**0.670**
-	si	0.110	0.062	2.890	0.181
VSM	si	0.116	0.066	2.621	0.305
Audio Sim	no	0.004	0.002	0.382	0.044

EBN: Entropy-based Novelty
ADiv: Aggregated Diversity

Music Recommendation (Conclusions)

Semantically Enriched Graph improves novelty and diversity → better explore the long tail

Combination with collaborative features ensures high accuracy

Path-based mapping: better novelty and diversity, slightly lower accuracy

Entity-based mapping: better accuracy, slightly lower novelty and diversity

Interpreting Music Recommendations

Building natural language **explanations** of the relation between two entities

- Using labels of a Knowledge Graph

  Fang, L., Sarma, A. A. Das, Yu, C., & Bohannon, P. (2011). REX: Explaining Relationships Between Entity Pairs. *Proceedings of the VLDB Endowment (PVLDB).*

  Passant, A. (2010). Dbrec—music recommendations using DBpedia. *The Semantic Web–ISWC 2010*, *1380*, 1–16.

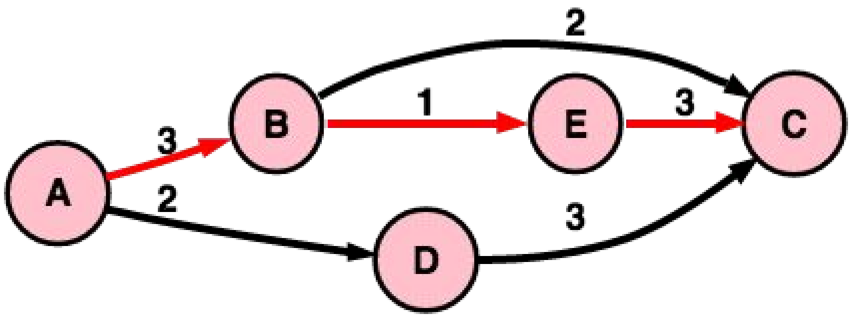
- Using sentence texts where entities co-occur

  Voskarides, N., & Meij, E. (2015). Learning to Explain Entity Relationships in Knowledge Graphs. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*, 564–574.

Interpreting Music Recommendations

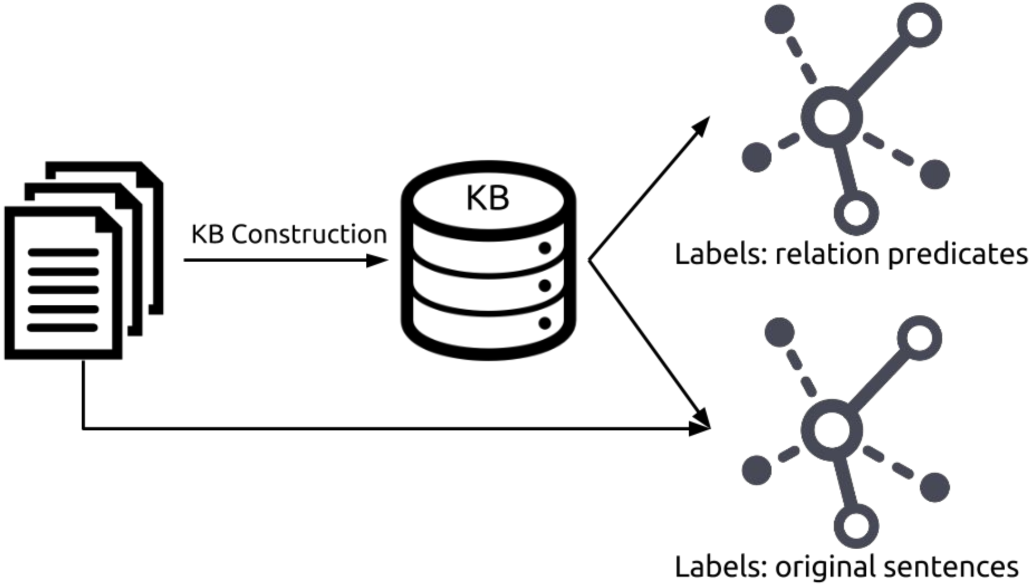
Challenges

- Select the best path (many possible paths between 2 entities)
- Generate a natural language explanation
  - Use relation labels
  - Use sentence texts



Interpreting Music Recommendations

Oramas S., Espinosa-Anke L., Sordo M., Saggion H., Serra X. (2016). Information Extraction for Knowledge Base Construction in the Music Domain. Journal on Knowledge & Data Engineering, Elsevier.



Interpreting Music Recommendations

User Experiment:

- 35 subjects
- 3 different recommendations
  - no explanation (3.08)
  - original sentences (**3.20**)
  - predicate labels (3.04)
- Higher differences in average ratings on musically untrained subjects

SONG #18

You Know My Name (Look Up The Number) (The Beatles)

RECOMMENDED SONG

Fourth Time Around (Bob Dylan)

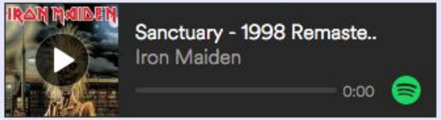
You Know My Name (Look Up The Number) <-- The Beatles <-- Fourth Time Around

The Beatles started recording You Know My Name (Look Up The Number) in 1967 , adding all the instrumentation and a saxophone part played by Brian Jones from The Rolling Stones . Fourth Time Around was written in response to `` Norwegian Wood -LRB- This Bird Has Flown -RRB- " by The Beatles , since it is similar , both melodically and lyrically .

Give a score to the provided recommendation:
○ 1 ○ 2 ○ 3 ○ 4 ○ 5
Did you know the recommended song?
○ Yes ○ No

SONG #10

Sanctuary (Iron Maiden)

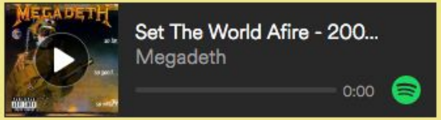


RECOMMENDED SONG

Set The World Afire (Megadeth)

Sanctuary <-- Iron Maiden <-- Jump In The Fire (Metallica) <-- Dave Mustaine --> Set The World Afire

Iron Maiden version of Sanctuary
Jump In The Fire (Metallica) was inspired by Iron Maiden
Dave Mustaine helped write Jump In The Fire (Metallica)
Dave Mustaine started writing Set The World Afire

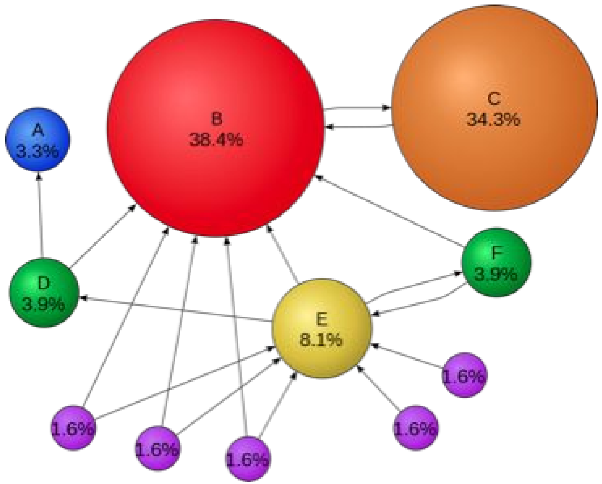


Give a score to the provided recommendation:
○ 1 ○ 2 ○ 3 ○ 4 ○ 5
Did you know the recommended song?
○ Yes ○ No

Artist Relevance

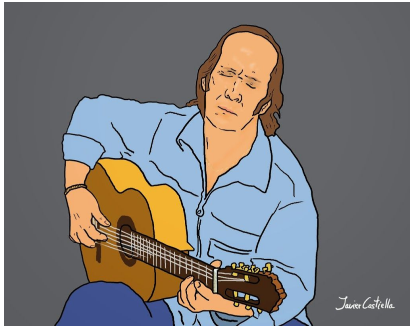
See a Graph of Entities as network of hyperlinks

Use Pagerank or HITS to compute entity relevance



Artist Relevance

Oramas S., Gómez F., Gómez E., & Mora J. (2015). FlaBase: Towards the creation of a Flamenco Music Knowledge Base. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).



Cantaor	Guitarist	*Bailaor*
Antonio Mairena	Paco de Lucía	Antonio Ruiz Soler
Manolo Caracol	Ramón Montoya	Rosario
La Niña de los Peines	Niño Ricardo	Antonio Gades
Antonio Chacón	Manolo Sanlúcar	Mario Maya
Camarón de la Isla	Sabicas	Carmen Amaya

Flamenco expert evaluation

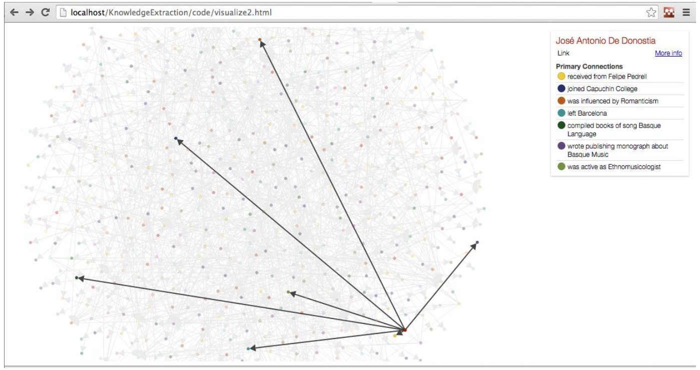
	Top-5	Top-10
PageRank	0.933	0.633
HITS Authority	0.6	0.4

Information Visualization

Extract a Knowledge Base from a Digital Library.

Build a Knowledge Graph to navigate through the library.

Create a visual representation of the graph.



Oramas S., Sordo M., & Serra X. (2014). Automatic Creation of Knowledge Graphs from Digital Musical Document Libraries. Conference in Interdisciplinary Musicology (CIM 2014)

Analytics

Extract attributes, events, entity mentions, relations.

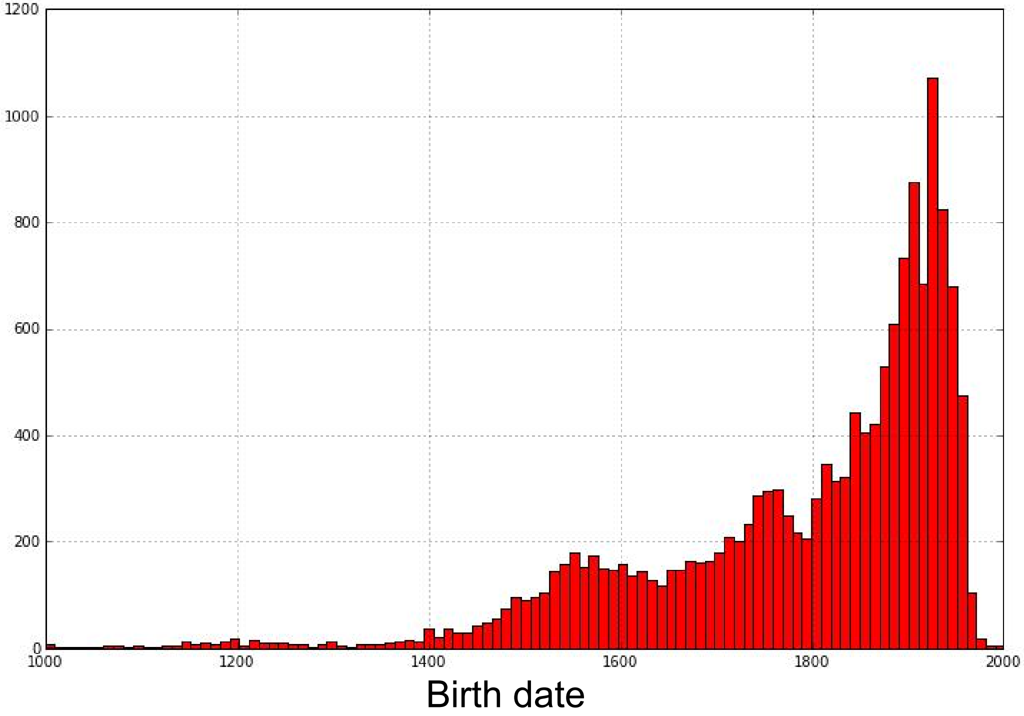
- Compute analytics, similarity, relevance
- Useful insights for musicologists

Analytics

Oramas S., Sordo M. (2016). Knowledge is Out There: A New Step in the Evolution of Music Digital Libraries. Fontes Artis Musicae, Vol 63, no. 4.

- 16,707 biographies from Grove Music Online
- Extracted roles, birth and death dates and places, entity mentions

Analytics



Role	Amount
composer	2618
teacher	1065
conductor	968
pianist	704
organist	676
singer	404
violinist	285
...	
musicologist	144
critic	133

Analytics

Country	Births	Deaths	Difference
United States	2317	2094	-10%
Italy	1616	1279	-21%
Germany	1270	1292	2%
France	991	1058	7%
United Kingdom	882	877	-1%

City	Births	Deaths	Difference
London	322	507	57%
Paris	304	720	137%
New York	266	501	88%
Vienna	177	292	65%
Rome	159	256	61%

Analytics

City	Births	Deaths	Difference
London	322	507	57%
Paris	**304**	**720**	**137%**
New York	266	501	88%
Vienna	177	292	65%
Rome	159	256	61%

Other Applications

- Question & Answering

Fader, A., Zettlemoyer, L., & Etzioni, O. (2014). Open question answering over curated and extracted knowledge bases. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '14*, 1156–1165.

- Entity Retrieval / Semantic Search

http://edgar.meij.pro/entity-linking-retrieval-semantic-search-wsdm-2014/

References

Oramas S., Sordo M., Espinosa-Anke L., & Serra X. (2015). *A Semantic-based approach for Artist Similarity*. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

Oramas S., Sordo M., Espinosa-Anke L., & Serra X. (2015). *A Semantic-based approach for Artist Similarity*. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

Fang, L., Sarma, A. A. Das, Yu, C., & Bohannon, P. (2011). REX: Explaining Relationships Between Entity Pairs. *Proceedings of the VLDB Endowment (PVLDB).*

Passant, A. (2010). Dbrec—music recommendations using DBpedia. *The Semantic Web–ISWC 2010*, *1380*, 1–16.

Voskarides, N., & Meij, E. (2015). Learning to Explain Entity Relationships in Knowledge Graphs. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*, 564–574.

Oramas S., Espinosa-Anke L., Sordo M., Saggion H., Serra X. (2016). Information Extraction for Knowledge Base Construction in the Music Domain. Journal on Knowledge & Data Engineering, Elsevier.Oramas S., Gómez F., Gómez E., &

References

Oramas S., Gómez F., Gómez E., & Mora J. (2015). FlaBase: Towards the creation of a Flamenco Music Knowledge Base. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

Oramas S., Sordo M., & Serra X. (2014). Automatic Creation of Knowledge Graphs from Digital Musical Document Libraries. Conference in Interdisciplinary Musicology (CIM 2014).

Oramas S., Sordo M. (2016). Knowledge is Out There: A New Step in the Evolution of Music Digital Libraries. Fontes Artis Musicae, Vol 63, no. 4.

Fader, A., Zettlemoyer, L., & Etzioni, O. (2014). Open question answering over curated and extracted knowledge bases. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '14*, 1156–1165.

Supplementary Material

Download supplementary material:

http://mtg.upf.edu/nlp-edu

Create a BabelNet account:

http://babelnet.org/register

Outline

- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - Applications in MIR
- **Topic Modeling**
- Sentiment Analysis
- Lexical Semantics

Topic Modeling

Topic modeling: motivations

- Many NLP tasks: NLP pipeline requires significant resources
  - annotated corpora
  - linguistic knowledge of the language
  - language specific NLP techniques
- Text mining and topic modeling: unsupervised approach without the above resources
- Design choices in MIR tasks

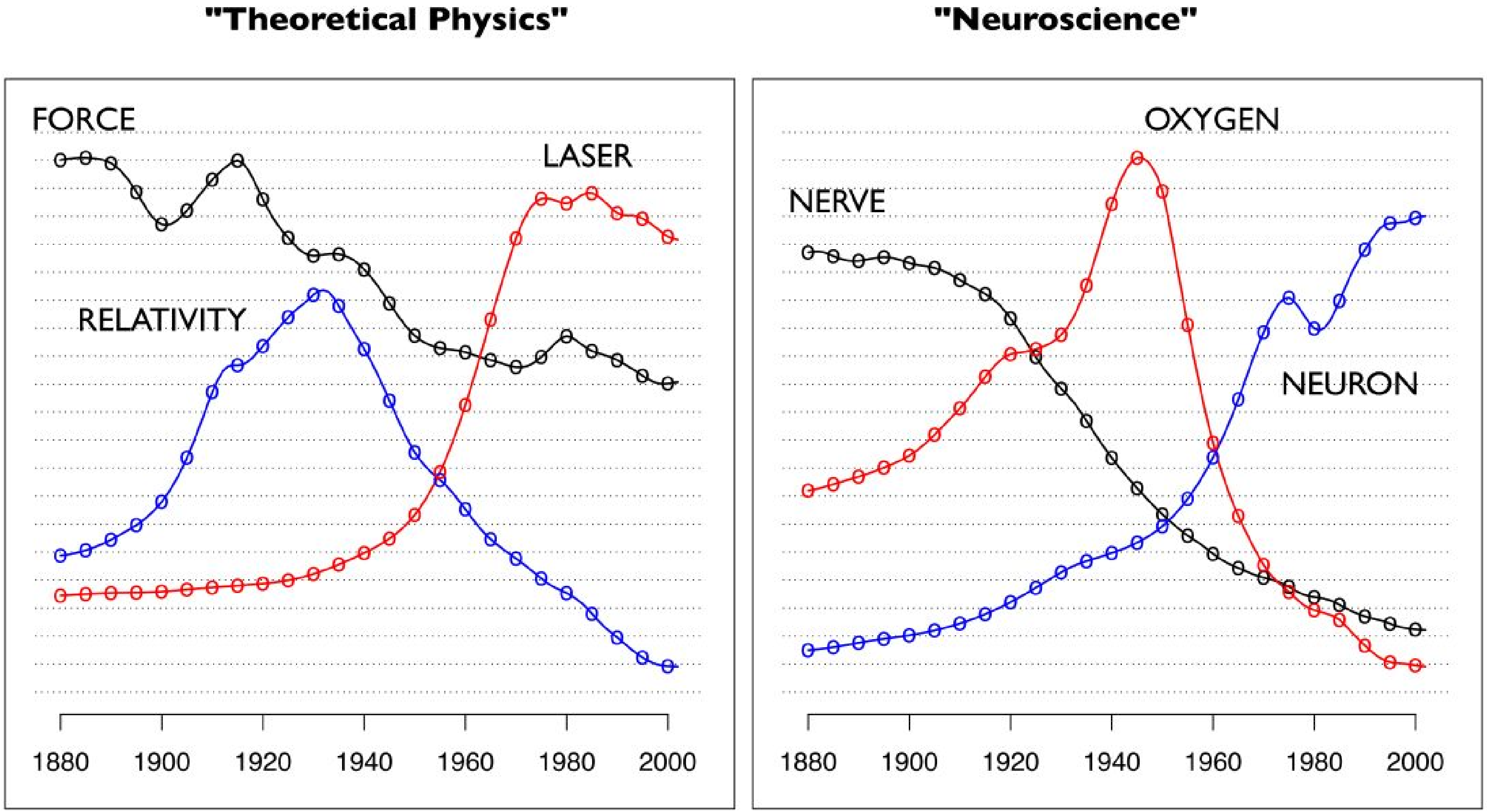
What is topic modeling?

- Unsupervised discovery of ‘abstract’ themes and topics that make up of a large body of text collections
- “Topic models are a suite of algorithms that uncover the hidden thematic structure in document collections. These algorithms help us develop new ways to search, browse and summarize large archives of texts.”

- reduced representation of collections of text documents

human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

Blei 2012



"Theoretical Physics"
FORCE
LASER
RELATIVITY
1880
1900
1920
1940
1960
1980
2000
"Neuroscience"
OXYGEN
NERVE
NEURON
1880
1900
1920
1940
1960
1980
2000

Blei 2012

topic	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14
MIR Data & fundamentals															
mus. signal processing	17.1	-	-	-	-	-	8.0	-	-	**19.5**	-	-	10.9	10.2	12.3
metadata, & semantic web	11.4	5.6	-	**17.0**	12.5	11.5	**16.1**	12.7	10.5	9.8	11.8	9.8	-	-	-
social tags & user gen. data	-	-	-	-	-	-	-	**13.5**	10.5	12.2	10.9	**12.8**	11.9	**12.2**	-
lyrics & genres & moods	-	-	-	-	-	-	-	-	11.4	11.4	-	10.5	9.9	-	11.3
Domain Knowledge															
comp. music. & ethnomus.	-	8.3	-	-	-	-	-	-	-	-	-	-	-	-	-
mus. notation	-	8.3	-	-	-	-	-	-	-	-	-	-	-	-	-
mir & cultures	-	-	-	-	-	-	-	-	-	-	-	9.8	-	10.2	-
Mus. Features & Properties															
melody & motives	11.4	-	11.3	8.5	8.7	-	9.2	11.9	-	-	-	11.3	12.9	-	-
harmony, chords & tonality	-	13.9	-	-	**13.5**	8.8	9.2	10.3	9.5	13.0	10.9	10.5	11.9	10.2	-
rhythm, beat, tempo	-	**19.4**	-	12.8	**13.5**	12.4	-	-	**13.3**	8.9	11.8	-	-	**12.2**	12.3
mus. affect, emot. & mood	-	-	-	10.6	-	-	-	-	-	-	-	-	-	10.2	-
structure, segment. & form	-	-	11.3	-	-	-	-	-	10.5	12.2	10.0	12.0	8.9	10.2	13.2
Music Processing															
sound source separation	-	-	-	-	-	-	8.0	10.3	-	-	**13.6**	-	**14.9**	**12.2**	11.3
mus. transcrip. & annot.	5.7	8.3	-	-	-	11.5	-	-	-	-	-	-	-	**12.2**	-
optical mus. recognition	-	-	-	-	-	-	6.9	10.3	-	-	-	-	9.9	-	-
align., synch. & score foll.	-	-	-	10.6	-	12.4	-	-	-	-	-	-	-	-	-
mus. summarization	-	-	7.5	-	-	-	-	-	-	-	-	-	-	-	-
fingerprinting	-	-	11.3	-	-	-	-	-	-	-	-	**12.8**	-	-	-
automatic classification	8.6	11.1	11.3	12.8	**13.5**	**14.2**	13.8	12.7	12.4	13.0	11.8	-	-	-	**14.2**
indexing & querying	**22.9**	13.9	9.4	10.6	7.7	9.7	9.2	-	-	-	10.9	-	-	-	-
pattern match. & detection	-	11.1	-	8.5	10.6	9.7	-	-	11.4	-	-	-	-	-	5.7
similarity metrics	-	-	-	8.5	9.6	-	11.5	8.7	-	-	8.2	10.5	-	-	-
Application															
user behavior & modeling	-	-	-	-	-	-	-	-	-	-	-	-	8.9	-	-
digital libraries & archives	11.4	-	-	-	10.6	-	-	-	-	-	-	-	-	-	-
mus. retrieval systems	-	-	**22.6**	-	-	-	-	-	10.5	-	-	-	-	-	8.5
mus. rec. & playlist gen.	-	-	15.1	-	-	9.7	8.0	-	-	-	-	-	-	-	11.3
mus. & gaming	-	-	-	-	-	-	-	9.5	-	-	-	-	-	-	-
mus. software	11.4	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Sordo et al., ISMIR15

Approaches to topic modeling

- Matrix factorization
- Probabilistic approaches
  - probabilistic Latent Semantic Indexing (pLSI)
  - Latent Dirichlet Allocation (LDA)

Vector representation of text documents

- A collection of n text documents $D_{1,}\ \dots, D_n$
- Vocabulary (Terms) $T_1,\dots,T_t$
- w_{ij} represents the weights of each term in the document (counts, binary, tf-idf, etc.)

- Sparse matrix with lots of zeros

- Bag of words (BOW) ignores word order

- “the department chair couches offers” vs. “the chair department offers couches”

$$
\begin{array}{c c}
 & \begin{array}{cccc} T_1 & T_2 & \dots & T_t \end{array} \\
\begin{array}{c} D_1 \\ D_2 \\ \vdots \\ \vdots \\ D_n \end{array} &
\left(\begin{array}{cccc}
w_{11} & w_{21} & \dots & w_{t1} \\
w_{12} & w_{22} & \dots & w_{t2} \\
\vdots & \vdots & & \vdots \\
\vdots & \vdots & & \vdots \\
w_{1n} & w_{2n} & \dots & w_{tn}
\end{array}\right)
\end{array}
$$

7

Topic modeling as matrix factorization problem

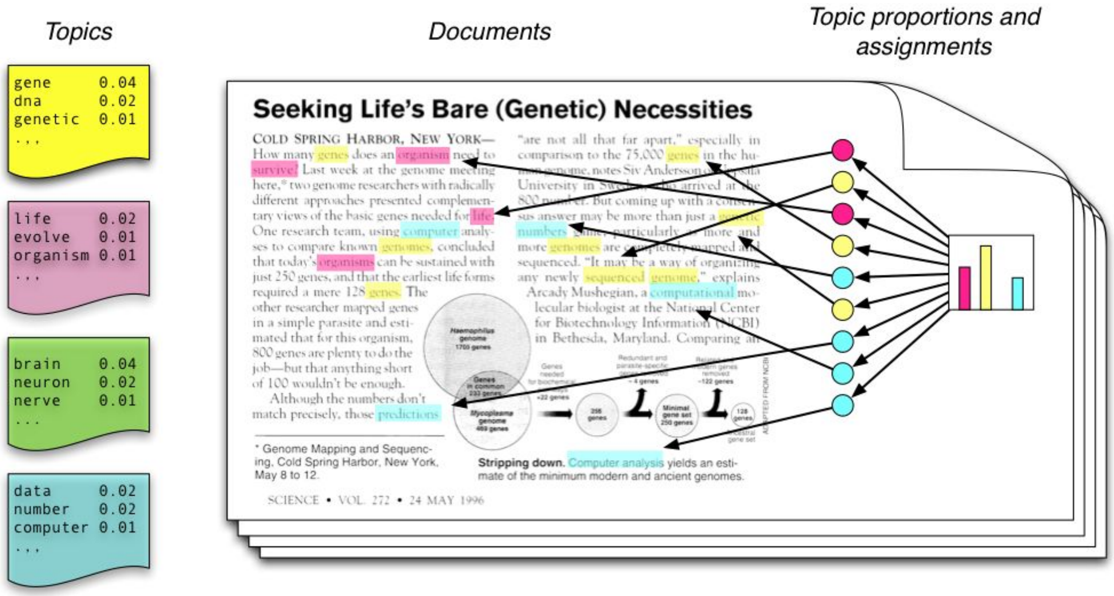
$$\underset{\text{Topic Assignment}}{\begin{bmatrix} M \times K \end{bmatrix}} \times \underset{\text{Topics}}{\begin{bmatrix} K \times V \end{bmatrix}} \approx \underset{\text{Dataset}}{\begin{bmatrix} M \times V \end{bmatrix}}$$

Boyd-Graber 2014

use SVD, NMF, etc.

- K Number of topics
- M Number of documents
- V Size of vocabulary

Probabilistic approaches to topic models



Blei 2012

generative story: words in text documents are generated by drawing from underlying topics.

- Each **topic** is a distribution over words
- Each **document** is a mixture of corpus-wide topics
- Each **word** is drawn from one of those topics

Probabilistic topic models:graphic view

Blei et al. 2003

$$p(\mathbf{w}) = \prod_{n=1}^{N} p(w_n).$$

w:word

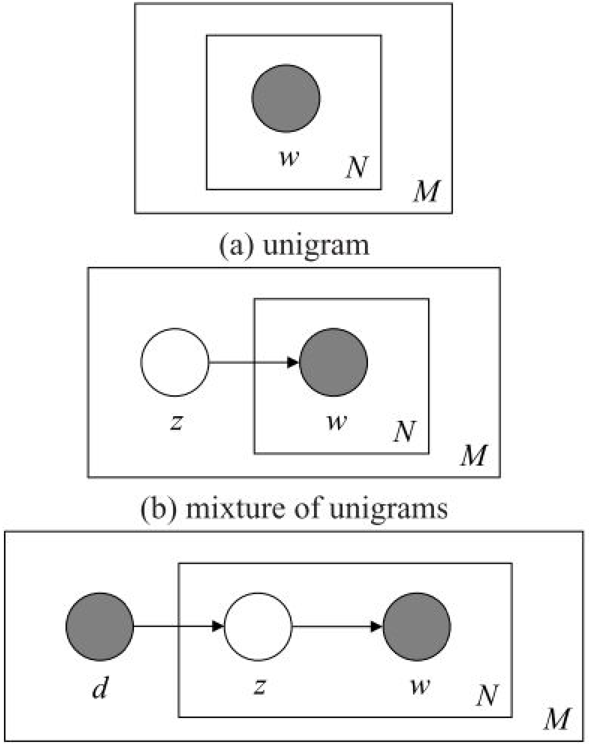
z:topic assignment

d: document

w: sequence of words

$$p(\mathbf{w}) = \sum_z p(z) \prod_{n=1}^{N} p(w_n \mid z).$$

$$p(d, w_n) = p(d) \sum_z p(w_n \mid z) p(z \mid d).$$

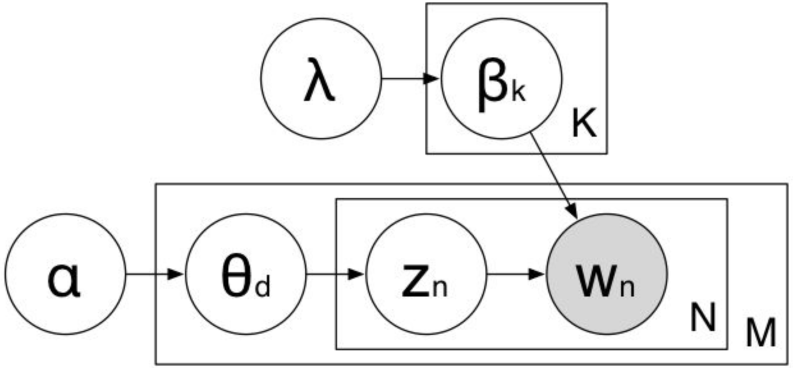


(a) unigram

(b) mixture of unigrams

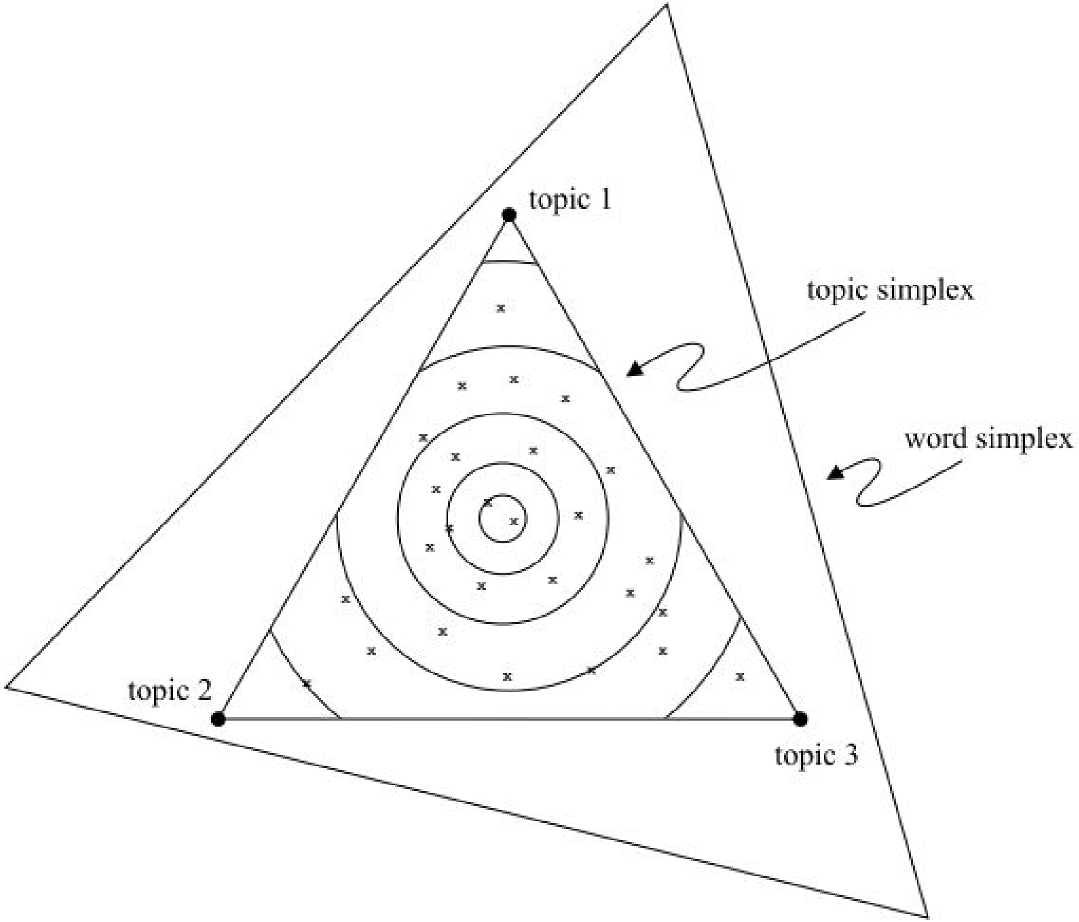
(c) pLSI/aspect model

Latent Dirichlet Allocation (LDA)



Boyd-Graber
2014

- For each topic $k \in \{1, \ldots, K\}$, draw a multinomial distribution β_k from a Dirichlet distribution with parameter λ
- For each document $d \in \{1, \ldots, M\}$, draw a multinomial distribution θ_d from a Dirichlet distribution with parameter α
- For each word position $n \in \{1, \ldots, N\}$, select a hidden topic z_n from the multinomial distribution parameterized by θ.
- Choose the observed word w_n from the distribution β_{z_n}.

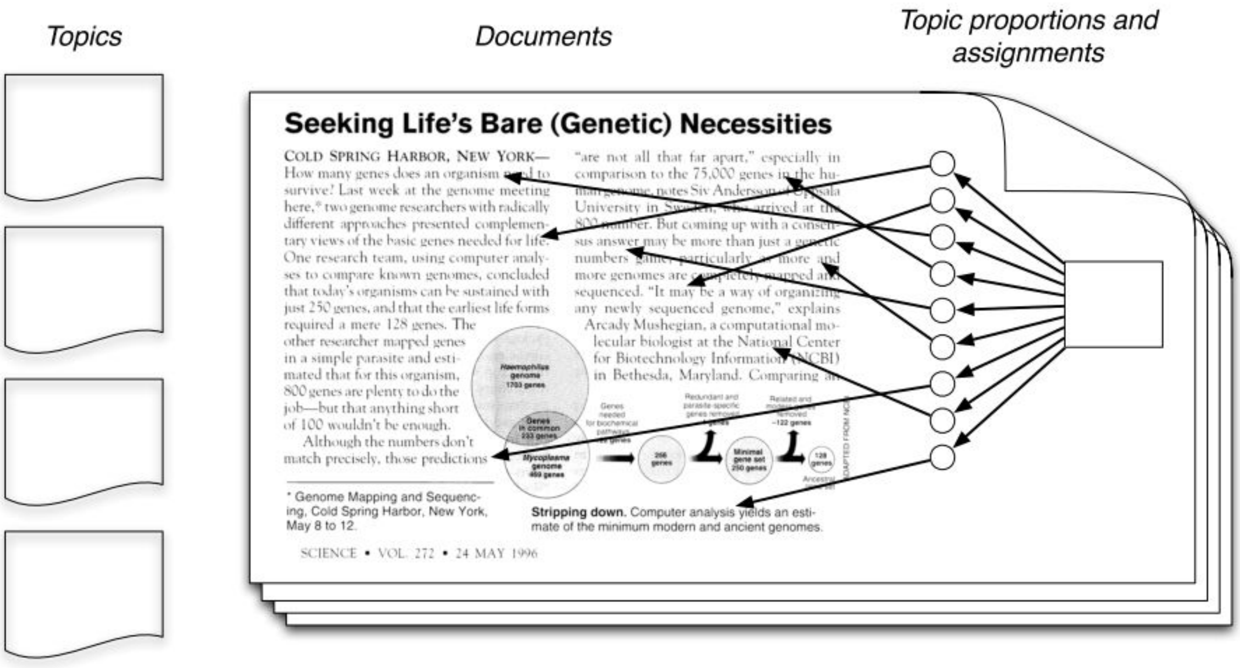


Blei et al 2003

geometric view of a topic model with 3 words and 3 topics

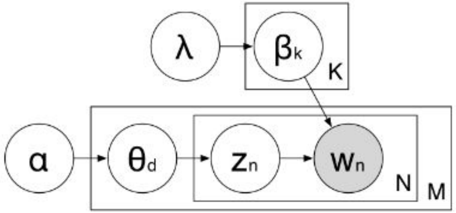
comparison of different probabilistic models

LDA



- In reality, we only observe the documents
- The other structure are **hidden variables**

LDA



- Topic to words: multinomial distribution with Dirichlet prior
- Document to topics: multinomial distribution with Dirichlet prior
- Joint distribution given by:

$$p(\mathbf{w}, \mathbf{z}, \boldsymbol{\theta}, \boldsymbol{\beta} | \alpha, \lambda) = \prod_k p(\beta_k|\lambda) \prod_d p(\theta_d|\alpha) \prod_n p(z_{d,n}|\theta_d) p(w_{d,n}|\beta_{z_{d,n}})$$

- Goal: posterior inference of hidden variables

$$p(\mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{w}, \alpha, \lambda)$$

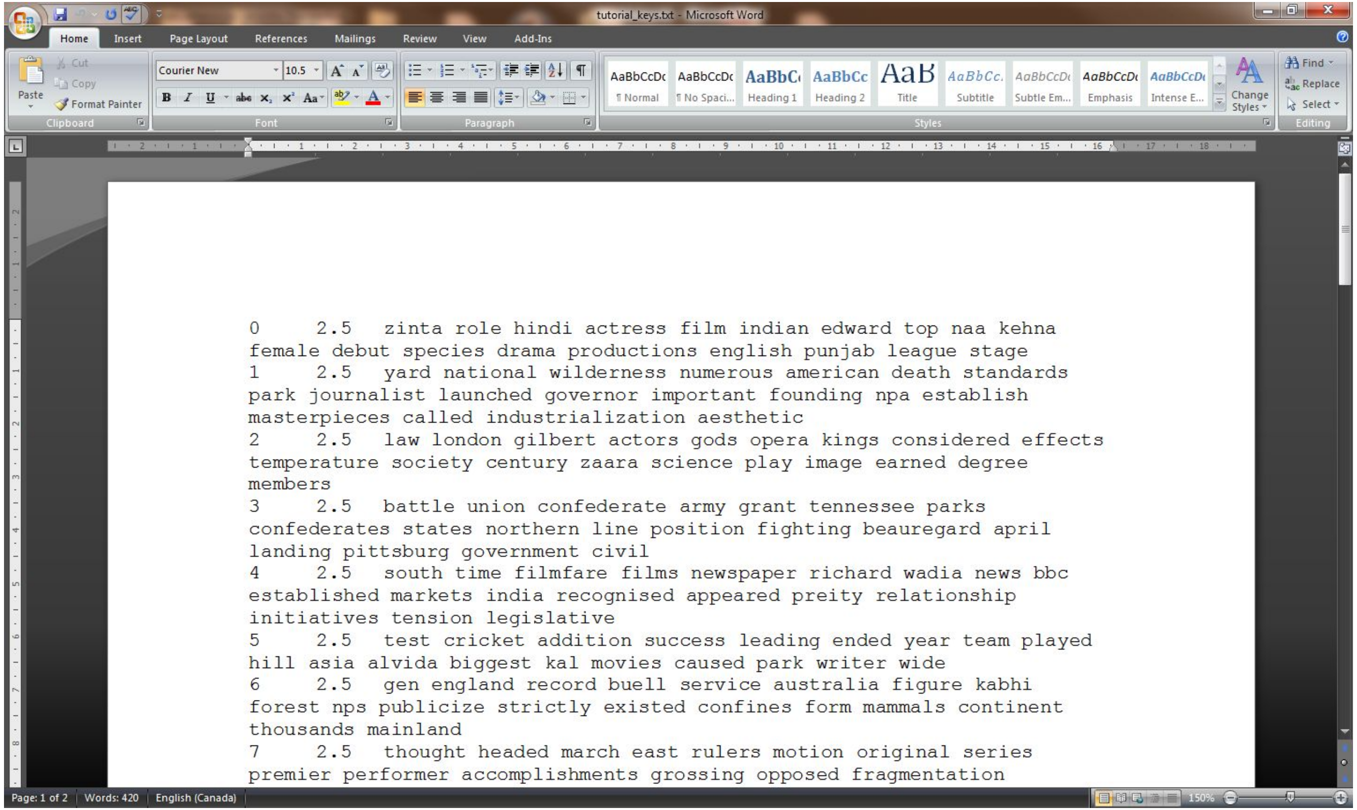
- Use MCMC, Gibbs sampling (sample topic assignment Z_n):

$$p(z_{d,n} = k | \mathbf{z}_{-d,n}, \mathbf{w}, \alpha, \lambda) = \frac{p(z_{d,n} = k, \mathbf{z}_{-d,n} | \mathbf{w}, \alpha, \lambda)}{p(\mathbf{z}_{-d,n} | \mathbf{w}, \alpha, \lambda)}$$

Implementations of LDA

- **Mallet (http://mallet.cs.umass.edu) (java)**
- LDAC (http://www.cs.princeton.edu/ blei/lda-c) (C)
- python LDA
- *not recommended:* python Gensim (according to Jordan Boyd-Graber 2014)

Example output from MALLET



0 2.5 zinta role hindi actress film indian edward top naa kehna female debut species drama productions english punjab league stage
1 2.5 yard national wilderness numerous american death standards park journalist launched governor important founding npa establish masterpieces called industrialization aesthetic
2 2.5 law london gilbert actors gods opera kings considered effects temperature society century zaara science play image earned degree members
3 2.5 battle union confederate army grant tennessee parks confederates states northern line position fighting beauregard april landing pittsburg government civil
4 2.5 south time filmfare films newspaper richard wadia news bbc established markets india recognised appeared preity relationship initiatives tension legislative
5 2.5 test cricket addition success leading ended year team played hill asia alvida biggest kal movies caused park writer wide
6 2.5 gen england record buell service australia figure kabhi forest nps publicize strictly existed confines form mammals continent thousands mainland
7 2.5 thought headed march east rulers motion original series premier performer accomplishments grossing opposed fragmentation

topic-word distributions

more on MALLET tutorial:

http://programminghistorian.org/lessons/topic-modeling-and-mallet

Example output from MALLET

tutorial_composition.txt - Microsoft Excel

B13 file:/C:/Mallet/sample-data/web/en/zinta.txt

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	#doc name topic proportion ...																	
2	0	file:/C:/Mallet/sample-data/web/en/elizabeth_needl	2	0.149038	8	0.110577	1	0.081731	18	0.0625	9	0.0625	7	0.0625	0	0.052885	4	0.043
3	1	file:/C:/Mallet/sample-data/web/en/equipartition_th	19	0.373288	11	0.058219	1	0.058219	13	0.05137	0	0.05137	18	0.044521	8	0.037671	4	0.037
4	2	file:/C:/Mallet/sample-data/web/en/gunnhild.txt	6	0.304511	8	0.12406	4	0.078947	3	0.06391	17	0.033835	16	0.033835	14	0.033835	9	0.033
5	3	file:/C:/Mallet/sample-data/web/en/hawes.txt	14	0.280645	5	0.170968	12	0.054839	7	0.054839	15	0.048387	13	0.041935	17	0.035484	16	0.035
6	4	file:/C:/Mallet/sample-data/web/en/hill.txt	10	0.305369	16	0.11745	3	0.083893	7	0.050336	19	0.036913	2	0.036913	0	0.036913	18	0.030
7	5	file:/C:/Mallet/sample-data/web/en/shiloh.txt	5	0.294872	16	0.074359	1	0.074359	9	0.058974	8	0.058974	18	0.053846	12	0.038462	11	0.038
8	6	file:/C:/Mallet/sample-data/web/en/sunderland_ech	15	0.311644	12	0.085616	9	0.071918	3	0.058219	11	0.044521	14	0.037671	10	0.037671	8	0.037
9	7	file:/C:/Mallet/sample-data/web/en/thespis.txt	7	0.12069	0	0.106897	8	0.093103	18	0.07931	17	0.058621	11	0.058621	14	0.051724	16	0.044
10	8	file:/C:/Mallet/sample-data/web/en/thylacine.txt	17	0.186111	1	0.108333	9	0.097222	3	0.086111	11	0.047222	15	0.041667	10	0.041667	8	0.041
11	9	file:/C:/Mallet/sample-data/web/en/uranus.txt	13	0.289809	18	0.092357	19	0.073248	11	0.05414	12	0.047771	9	0.047771	4	0.047771	0	0.047
12	10	file:/C:/Mallet/sample-data/web/en/yard.txt	12	0.34472	16	0.059006	0	0.052795	15	0.046584	11	0.046584	7	0.046584	10	0.040373	5	0.040
13	11	file:/C:/Mallet/sample-data/web/en/zinta.txt	2	0.230159	10	0.119048	7	0.07672	4	0.07672	17	0.060847	8	0.055556	12	0.050265	16	0.044

tutorial_composition

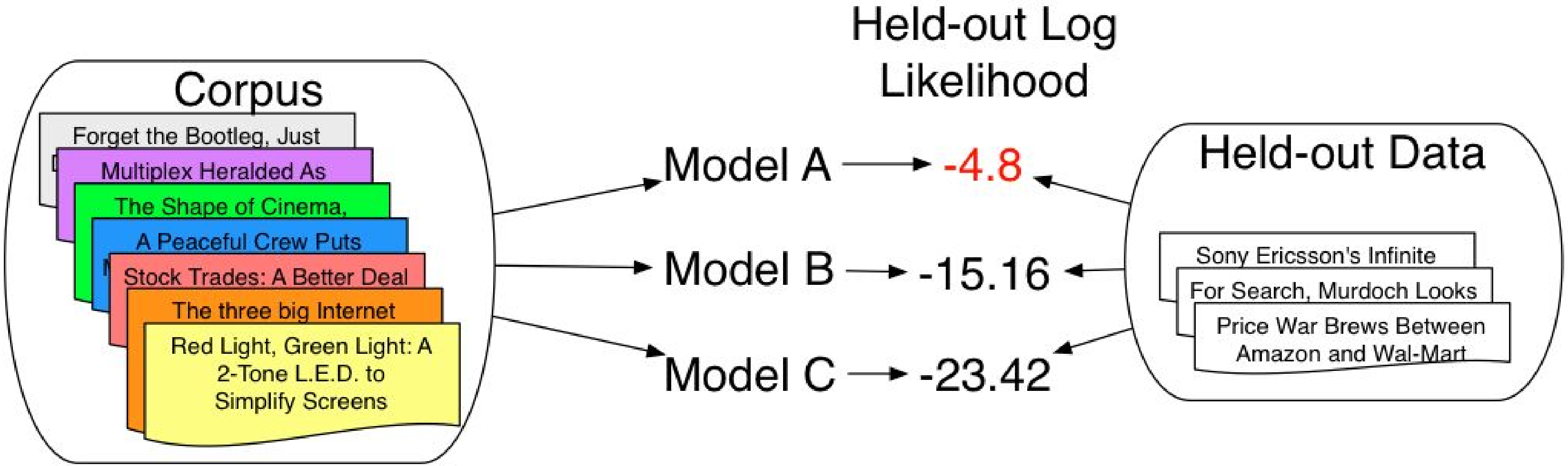
Document-topic distributions

LDA:evaluation and considerations

- How to choose number of latent topics: compute perplexity on a held out data set (Blei et al. 2003)
- Conventionally used in language modeling, perplexity is monotonically decreasing in the likelihood of the test data
- Indicates how well a language model (or a topic model in this case) generalizes to unseen data
- Also shows why LDA is superior to simpler models like unigram and pLSI

$$perplexity(D_{\text{test}}) = \exp\left\{-\frac{\sum_{d=1}^{M} \log p(\mathbf{w}_d)}{\sum_{d=1}^{M} N_d}\right\}.$$

Evaluation:number of latent topics



Perplexity vs. human judgment(example:word intrusion task)

evaluating topic coherence

1. Take the highest probability words from a topic

Original Topic

dog, cat, horse, pig, cow

2. Take a high-probability word from another topic and add it

Topic with Intruder

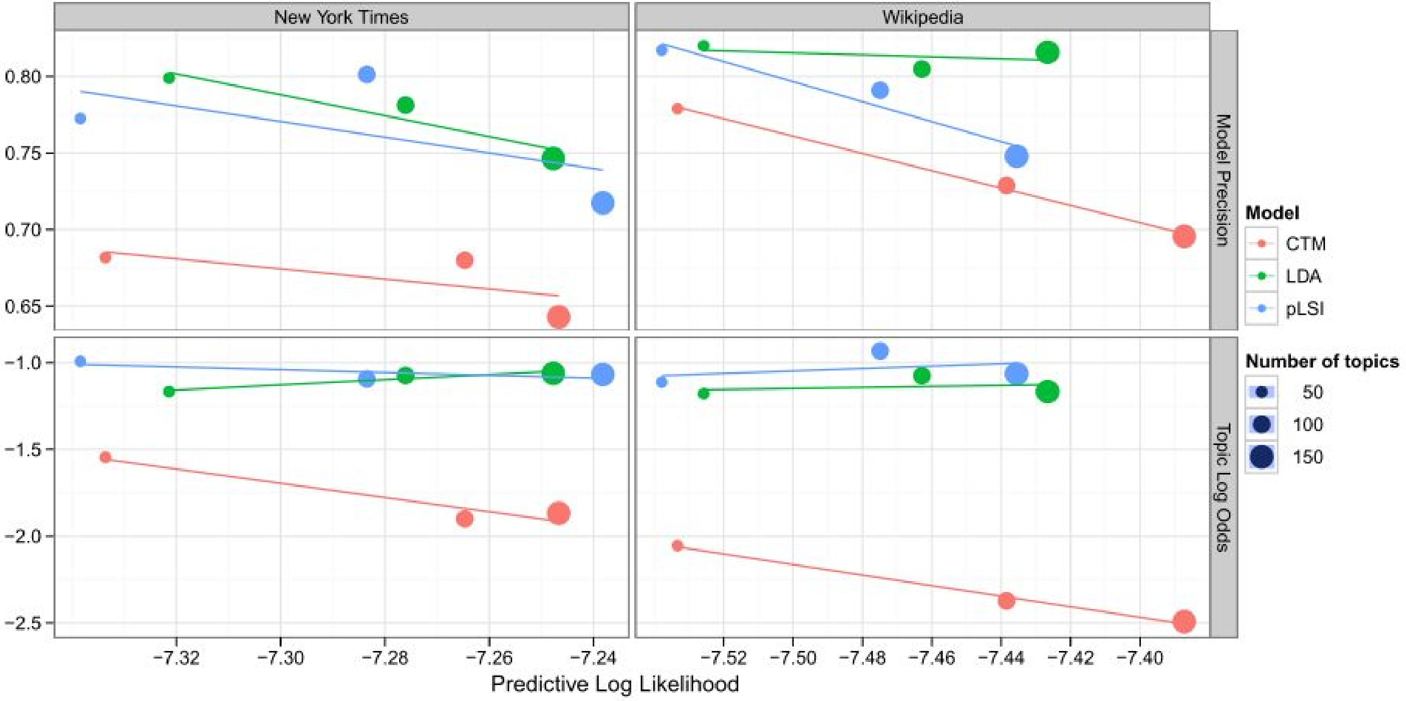
dog, cat, apple, horse, pig, cow

3. We ask users to find the word that doesn't belong

Hypothesis

If the topics are interpretable, users will consistently choose true intruder

Perplexity ≠ topic interpretabiliby!

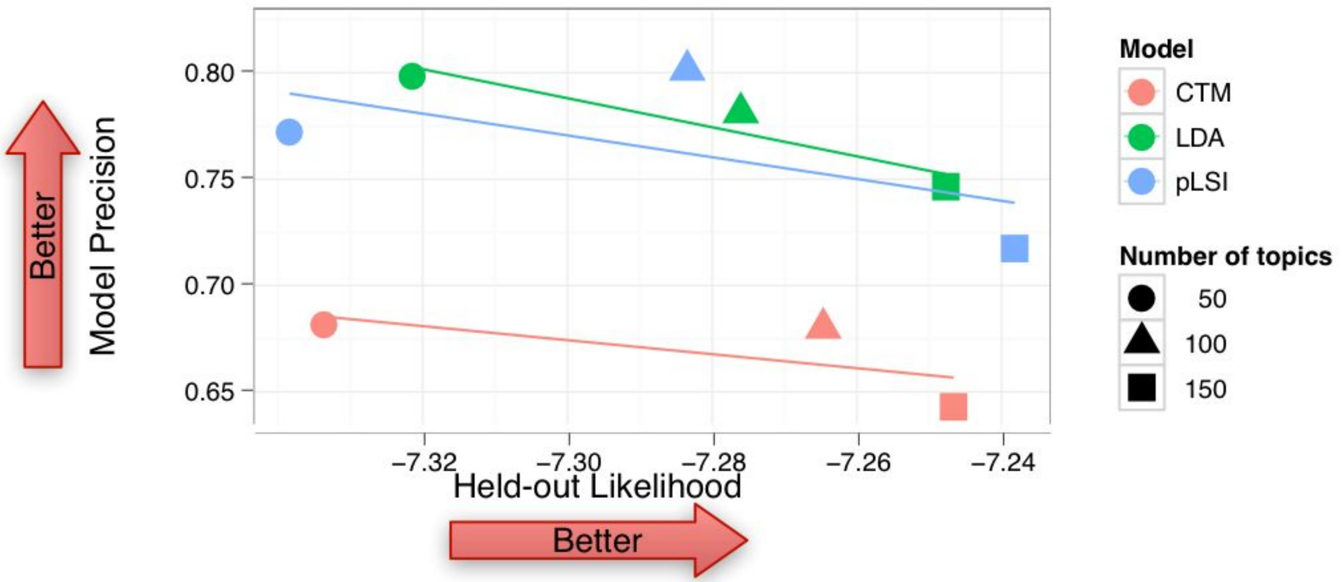


Chang et al. 2009

Model precision = topic coherence by human judgement

The assumption that latent topic automatically makes semantic sense is false.

Model Precision on New York Times

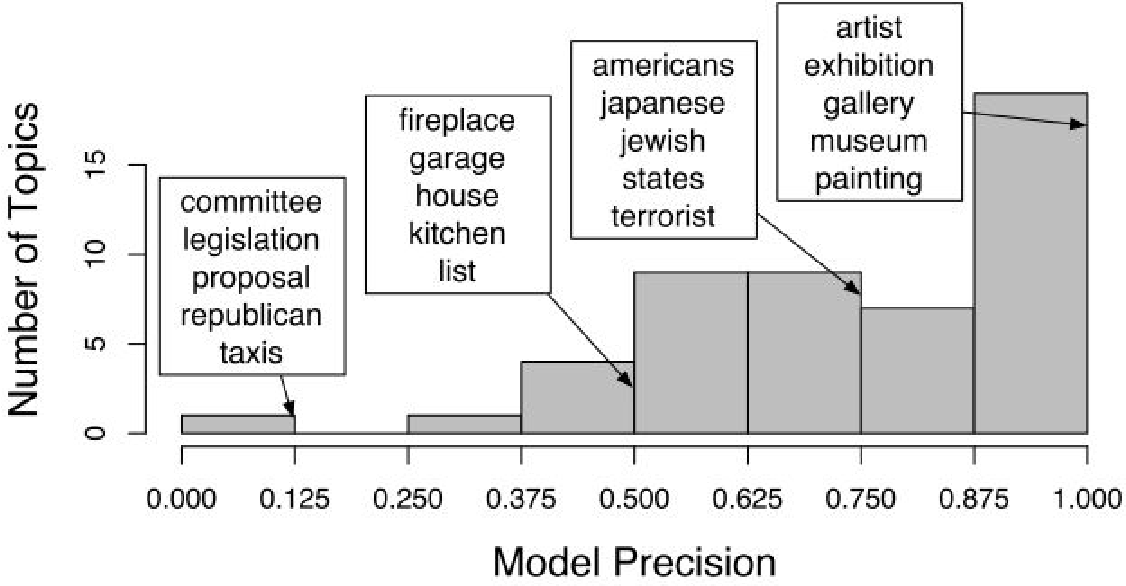


Boyd-Graber 2014

Lesson: measure what you care about

within a model, higher likelihood $\neq$ higher interpretability

Word intrusion task: example low and high quality topics(interpretability by human)



Chang et al. 2009

Histogram of model precisions on the NYT corpus with example good (high precision) and bad topics

More on evaluating topic coherence/quality

- Other automated evaluation on topic coherence
- normalized Pairwise Mutual Information (nPMI) (Alestras and Stevenson 2013; Lau et al. 2014)
- automated word intrusion task (Lau et al. 2014)
- Cardinality of topics: aggregated based on different values of *N* (Lau et al. 2016), where N=cardinality of topics, i.e., number of top ranked words in a topic distribution under evaluation

Domain	N	In-domain Features	Out-of-Domain Features
WIKI	5	0.46	0.66
	10	0.41	0.54*
	15	0.32*	0.51*
	20	0.33*	0.43*
	Avg	0.46	0.65
NEWS	5	0.45*	0.65
	10	0.40*	0.60*
	15	0.38*	0.54*
	20	0.43*	0.47*
	Avg	0.50	0.65

Table 3: Pearson correlation between system model precision and human ratings across different values of N for word intrusion. '*' denotes statistical significance compared to aggregate correlation.

Lau et al. 2014

N=cardinality of topics

Other considerations with LDA (and topic models in general)

- Short documents
- Dynamic Topic Modeling (modeling influence over time)
- Linguistic extensions of topic models (Boyd-Graber 2010)
- n-gram language modeling with LDA (vs. BOW repr.)
- automatically learn the number of topics(Blei et al. 2010)

Topic modeling in MIR examples

Hu et al ISMIR 2009	Using LDA to infer music key-profile - symbolic music files play the role of text documents, groups of musical notes play the role of words, and musical key-profiles play the role of topics
Sasaki et al ISMIR 2014	Lyrics Radar, a topic based music lyrics browser and exploration tool
Shalit et al 2013	Modeling chronic music influences using techniques similar to Dynamic topic modeling to discover influential scientific papers
Sordo et al ISMIR 2015	Using LDA to discover the evolution of research topics in past ISMIR proceedings over the years
Sterckx 2014	Assessing the quality of topics discovered in lyrics databases using LDA

Topic modeling in computational musicology example: Beijing opera at CompMusic Project

- Understanding the expressive functions of banshi (Rhythmic types) in Beijing opera using lyrics text mining and topic modeling
- **Premise**: Different *banshi* and *shengqiang*(melodic types) are used to express different themes, sentiment and expressive functions in Beijing opera music
- **Goal:** discover these functions through lyrics topics/sentiments associated with different *banshi* or *banshi-shengqiang* types
- **Problems**:
  - This research is exploratory
  - non-standard Chinese tokenization/NLP
  - Create data sets by fragmenting the lyrics corpus resulting in short documents

References

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet allocation. Journal of Machine Learning Research, 3:993–1022.

Jonathan Chang, Sean Gerrish, Chong Wang, Jordan L. Boyd-Graber, and David M. Blei. 2009. Reading tea leaves: How humans interpret topic models. In Advances in Neural Information Processing Systems 21 (NIPS-09), pages 288–296, Vancouver, Canada.

Jey Han Lau, David Newman, and Timothy Baldwin. 2014. Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In Proceedings of the 14th Conference of the EACL (EACL 2014), pages 530–539, Gothenburg, Sweden.

Nikos Aletras and Mark Stevenson. 2013. Evaluating topic coherence using distributional semantics. In Proceedings of the Tenth International Workshop on Computational Semantics (IWCS-10), pages 13–22, Potsdam, Germany.

Jordan Boyd-Graber. Topic Modeling (slides). From Natural Language Processing course taught at University of Colorado, Fall 2014 (http://www.umiacs.umd.edu/~jbg/teaching/CSCI_5832/1103.pdf).

Shoto Sasaki, Kazuyoshi Yoshii, Tomoyasu Nakano, Masataka Goto, and Shigeo Morishima. Lyricsradar: a Lyrics Retrieval System. ISMIR 2014:Proceedings of the 15th International Society for Music Information Retrieval Conference, (Ismir):585–590, 2014.

Uri Shalit, Daphna Weinshall, and Gal Chechik. Modeling Musical Influence with Topic Models. Proceedings of the 30th International Conference on Machine Learning (ICML), 28:244–252, 2013.

Mohamed Sordo, Mitsunori Ogihara, and Stefan Wuchty. Analysis of the Evolution of Research Groups and Topics in the Ismir Conference. 16th International Society for Music Information Retrieval Conference (ISMIR 2015), (Table 1):204–210, 2015.

L Sterckx, T Demeester, J Deleu, L Mertens, and C Develder. Assessing quality of unsupervised topics in song lyrics. Advances in Information Retrieval. 36th European Conference on IR Research, ECIR 2014. Proceedings: LNCS 8416, pages 547–552, 2014.

Jey Han Lau and Timothy Baldwin. The Sensitivity of Topic Coherence Evaluation to Topic Cardinality. Naacl, pages 483–487, 2016.

References

Vivek Kumar and Rangarajan Sridhar. Unsupervised Topic Modeling for Short Texts Using Distributed Representations of Words. Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing, pages 192–200, 2015.

Sean M Gerrish. A Language-based Approach to Measuring Scholarly Impact. Computer, 180(33):375–382, 2010.

Diane J Hu and Lawrence K Saul. A probabilistic topic model for unsupervised learning of musical key-profiles. Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009), (Ismir):441–446, 2009.

Jordan Boyd-graber. Linguistic Extensions of Topic Models. PhD Dissertation. 2010.

Blei, D. 2012. Introduction to probabilistic topic models. ICML tutorial slides, from http://www.cs.columbia.edu/~blei/topicmodeling.html.

Wallach, H. 2005. Topic modeling: beyond bag-of-words. In NIPS 2005.

Xuerui Wang, Andrew McCallum, Xing Wei. Topical N-grams: Phrase and Topic Discovery, with an Application to Information Retrieval. ICDM 2007.

D.Blei, T.Griffiths, and M.Jordan. The nested Chinese restaurant process and Bayesian nonparametric inference of topic hierarchies. Journal of the ACM, 57(2):1-30,2010.

Outline

- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - Applications in MIR
- Topic Modeling
- **Sentiment Analysis**
- Lexical Semantics

Sentiment Analysis

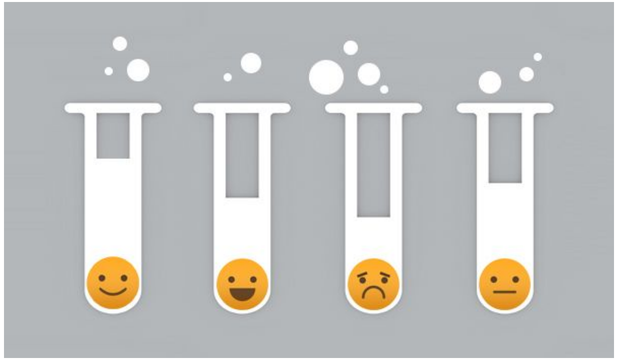
Sentiment Analysis

Computational study of opinions, sentiments, subjectivity, evaluations, attitudes, appraisal, affects, views, emotions, etc., expressed in text.

Complex NLP task

Pang, B., & Lee, L. (2006). *Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, 1(2). 91–231.*

https://www.cs.uic.edu/~liub/FBS/Sentiment-Analysis-tutorial-AAAI-2011.pdf



Document Sentiment Classification

Classify a whole opinion document (e.g., a review) based on the overall sentiment of the opinion holder. (positive, negative, neutral)

Great CD! I am nineteen years old and dont generally listen to piano music but I got this as a gift and I love it! The music is beautiful but also fun to listen to because the songs are familiar to me.

I don't care how much you hype this... it sucks...waste my money on this...never again!

Approaches:

- Unsupervised (polarity lexicons, heuristics)
- Supervised learning

Other Sentiment Analysis tasks

- Sentiment classification at sentence level
- Irony detection
- Aspect-based sentiment analysis
- Aspect-based opinion summarization

Other Sentiment Analysis tasks

- Sentiment classification at sentence level
- Irony detection
- **Aspect-based sentiment analysis**
- Aspect-based opinion summarization

Aspect-based Sentiment Analysis

Pos. Neg.

Beautiful Drug has great guitar riffs but the vocals are shrill

- Entities: Beautiful Drug
- Aspects (also called features): guitar riffs, vocals
- Opinion words: great, shrill

Tata, S., & Di Eugenio, B. (2010). Generating Fine-Grained Reviews of Songs from Album Reviews. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, (July), 1376–1385.

Ruihai Dong, Michael P O'Mahony, and Barry Smyth (2014). Further Experiments in Opinionated Product Recommendation. In ICCBR'14, pages 110–124.

Aspect-based Sentiment Analysis

Oramas S., Espinosa-Anke L., Lawlor A., Serra X., Saggion H. (2016). *Exploring Music Reviews for Music Genre Classification and Evolutionary Studies*. 17th International Society for Music Information Retrieval Conference. ISMIR 2016.

Rule-based approach using a sentiment lexicon

- Identification of aspects: bi-grams and single-noun
- Identification of opinion words: adjectives
- Context rules: distance, POS tags and negations between opinion words and aspects
- Sentiment Lexicon: SentiWordNet (http://sentiwordnet.isti.cnr.it/)

Exploring Music Reviews

MARD (Multimodal Album Reviews Dataset):

New dataset of album customer reviews from:

Amazon + MusicBrainz + AcousticBrainz

Experiments:

Album genre classification comparing semantic, sentiment and acoustic features

Diachronic study of affective language

Poster Session on Monday!

MARD Multimodal Album Reviews Dataset

- Genre annotations
- Amazon (66,566 albums / 263,525 reviews)
    - Album customer reviews
    - Genre tags (16 genres and 287 subgenres)
    - Star Ratings
    - Metadata: title, artist, record label
- MusicBrainz (28,053 albums): ids, song titles, year of publication
- AcousticBrainz (8,683 albums / 65,786 songs): audio descriptors of songs

MARD: http://mtg.upf.edu/download/datasets/mard

Diachronic Study of Affective Language

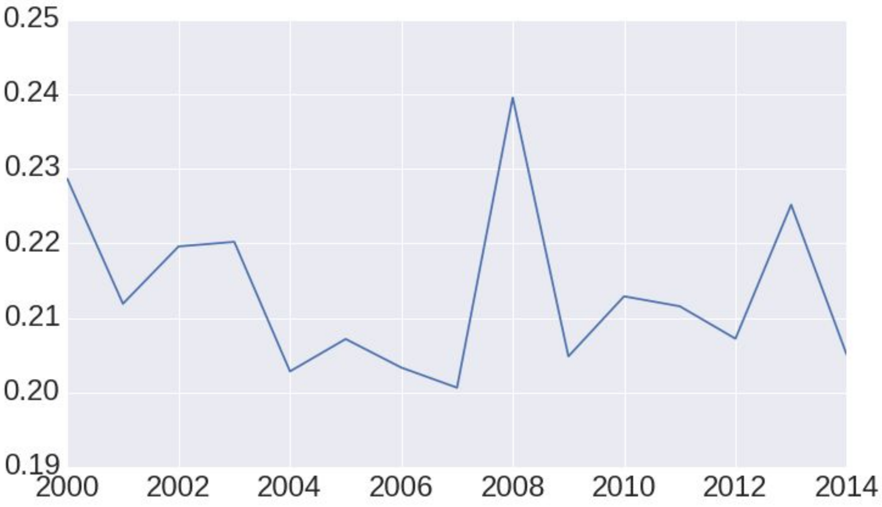
Sentiment score: Average sentiment score of all aspects in a review

Two perspectives:

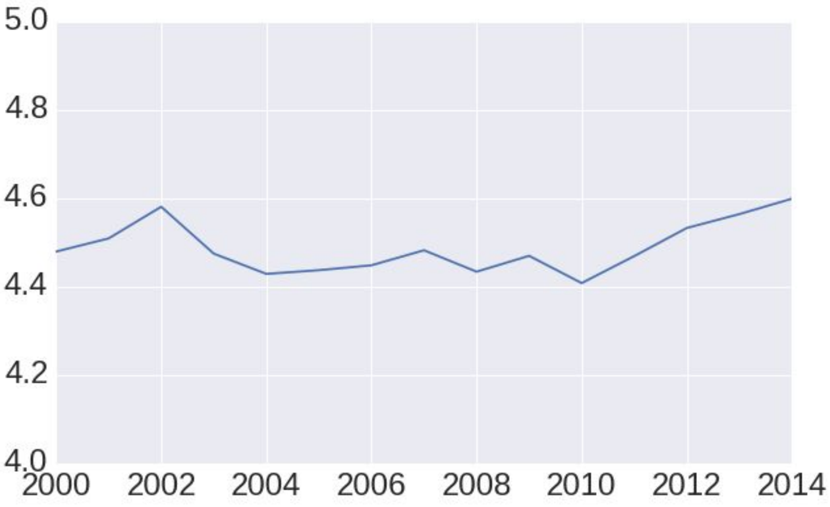
- Average of all reviews by review publication year (2000-2014)
  - Evolution of affective language from a customer perspective
- Average of all reviews by album publication year (1950-2014)
  - Evolution of affective language from a musical perspective

Study by review publication year

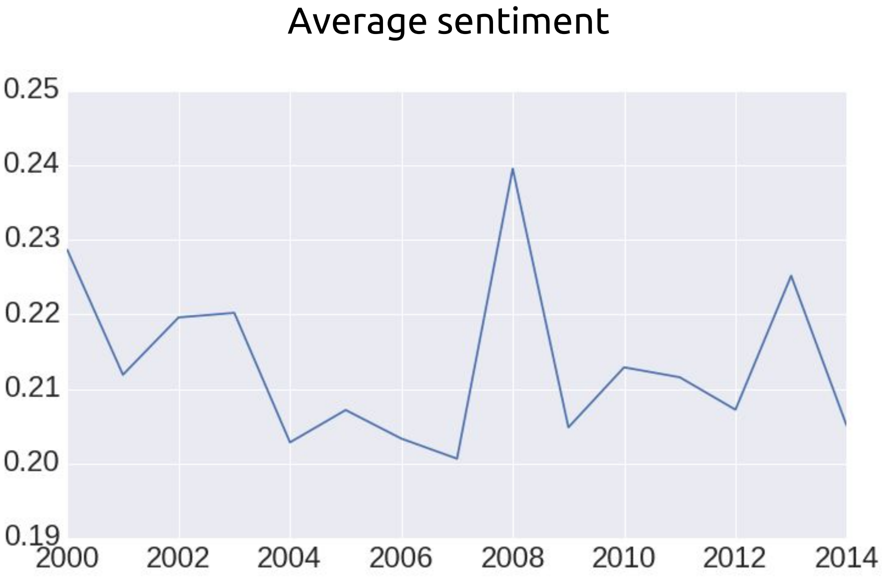
Average sentiment

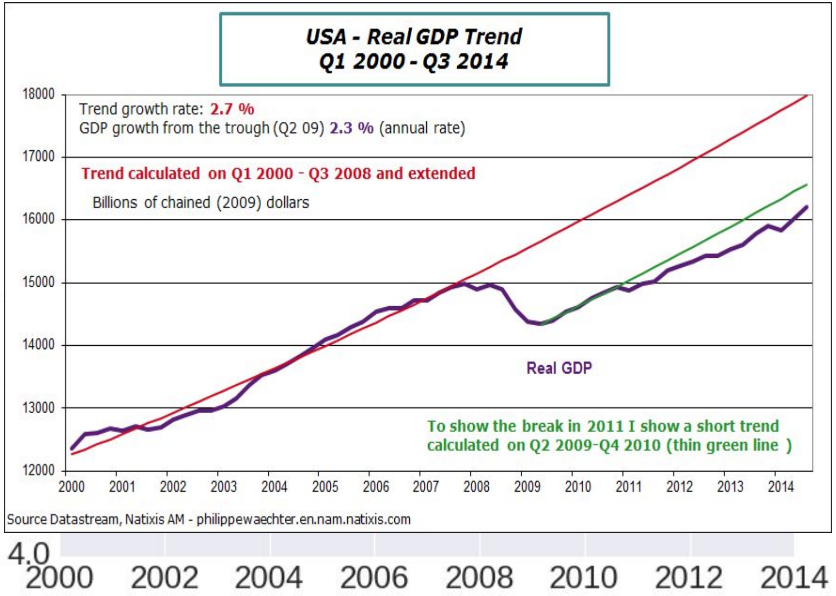


Average rating



Study by review publication year





Study by review publication year

Dominique Moïsi in:

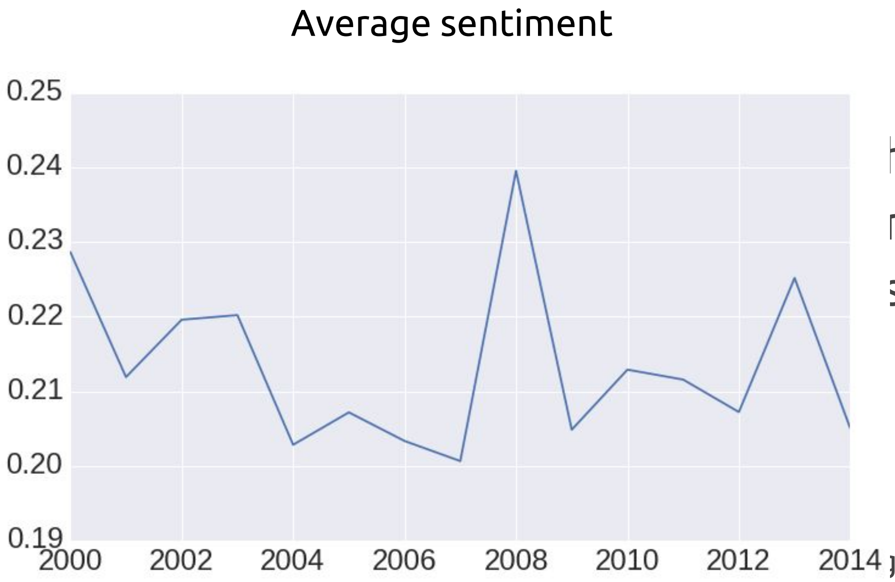
In November 2008, at least for a time, hope prevailed over fear. The wall of racial prejudice fell as surely as the wall of oppression had fallen in Berlin twenty years earlier [...] Yet the emotional dimension of this election and the sense of pride it created in many Americans must not be underestimated.

Dominique Moisi. The Geopolitics of Emotion: How Cultures of Fear, Humiliation, and Hope are Reshaping the World. Anchor Books, New York, NY, USA, 2010.

Study by review publication year

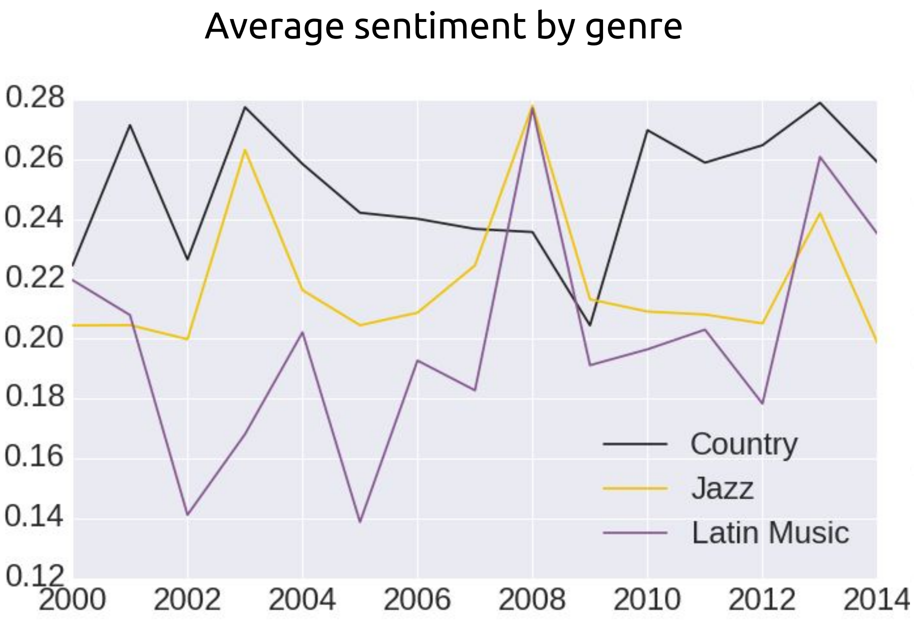
Dominique Moïsi in:

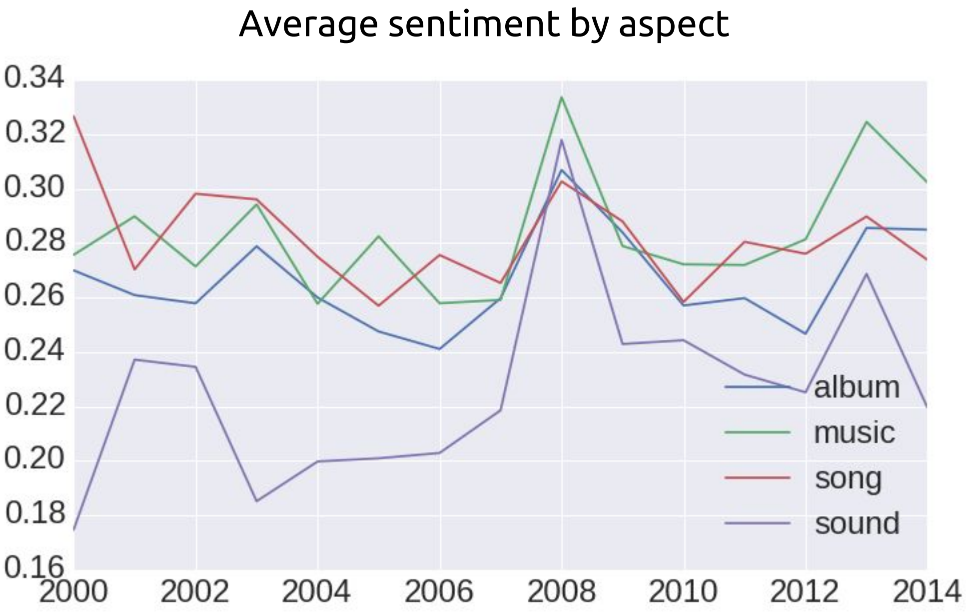
In November 2008, at lea he wall of racial prejudice fell as surely rlin twenty years earlier [...] Yet the emo sense of pride it created in many Ame



Dominique Moisi. The Geopolitics of Emotic the World. Anchor Books, New York, NY, USA, 2010.

Study by review publication year





Study by review publication year

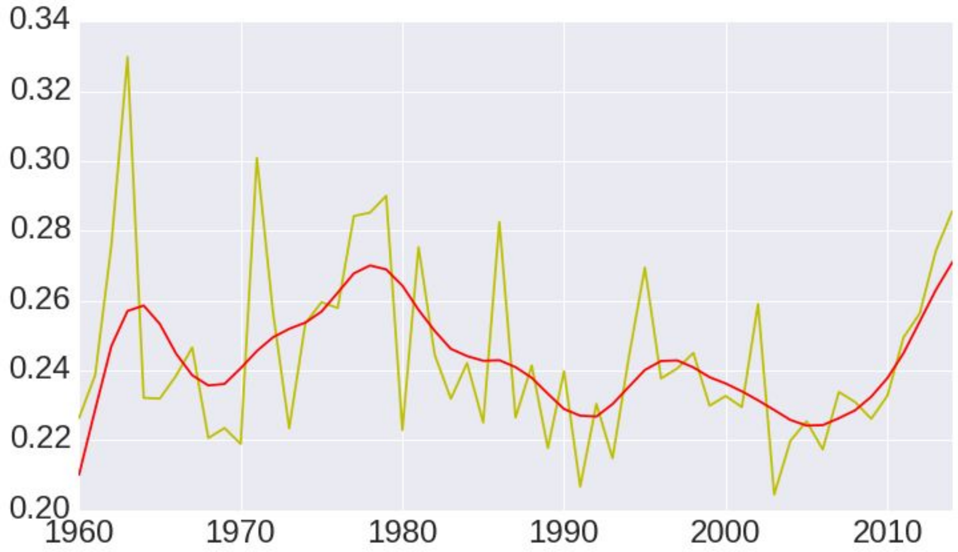
Further studies necessary to validate any of this suggestions

Correlation ≠ Causation

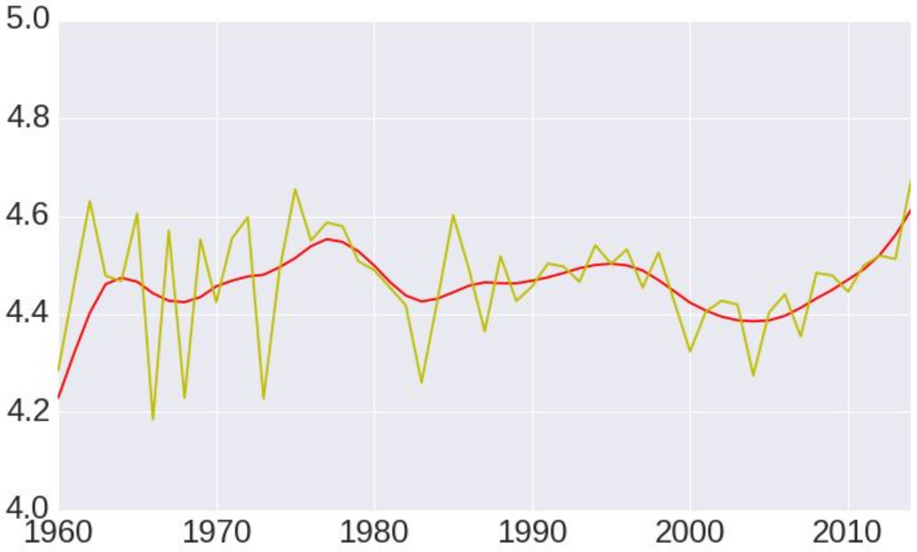
Interesting insight for Musicologists

Study by album publication year

Average sentiment



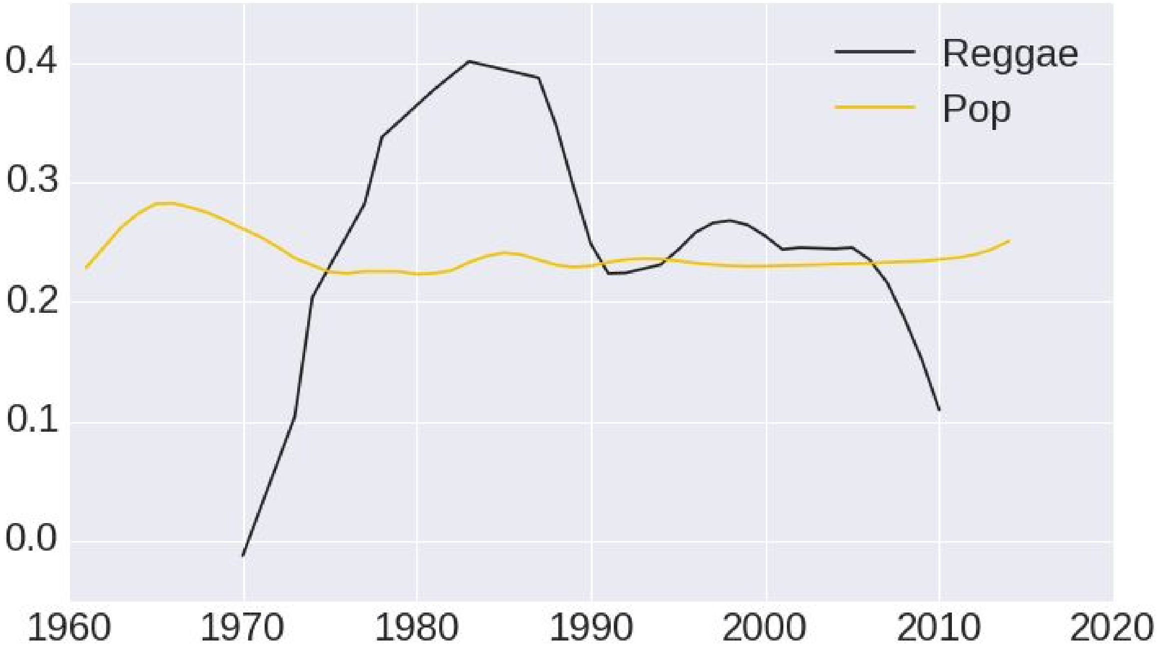
Average rating



Pearson's correlation r = 0.75, $p \ll 0.001$

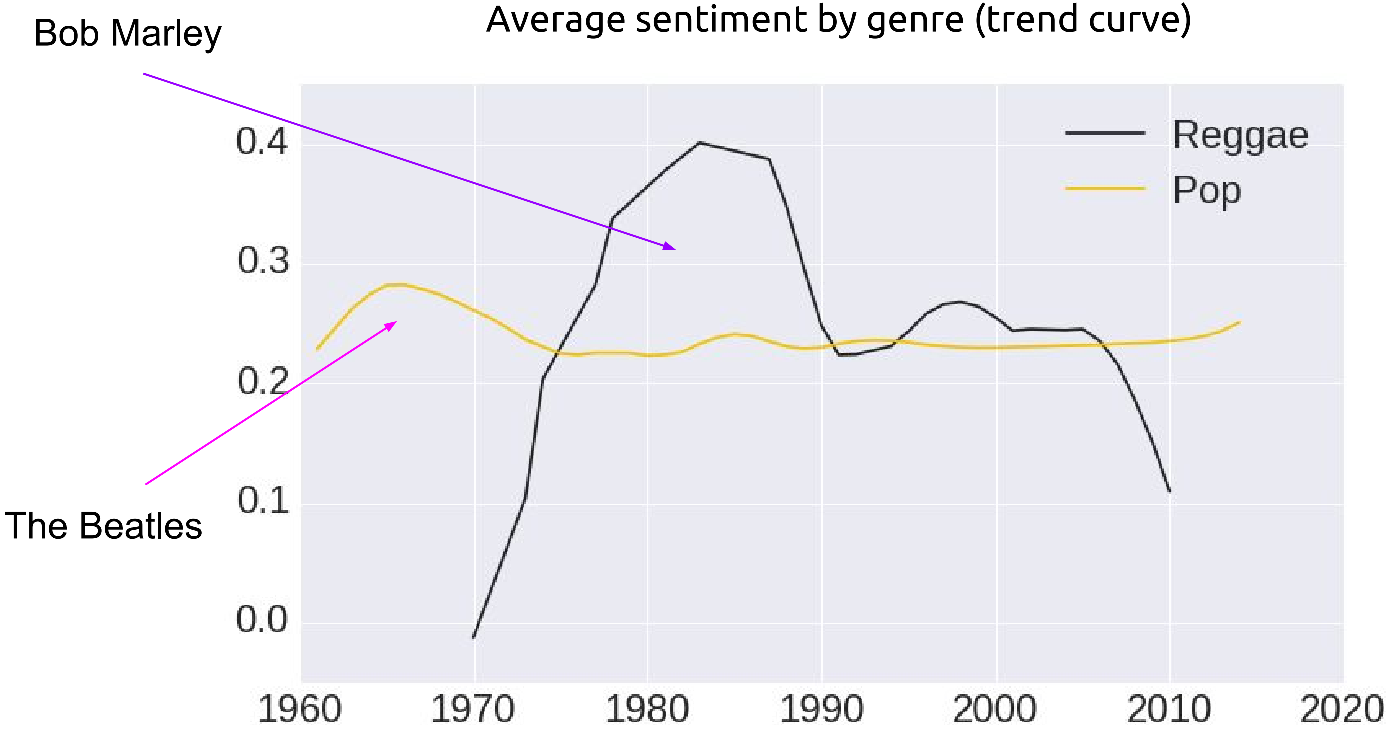
Study by album publication year

Average sentiment by genre (trend curve)



Study by album publication year

Average sentiment by genre (trend curve)



Study by album publication year

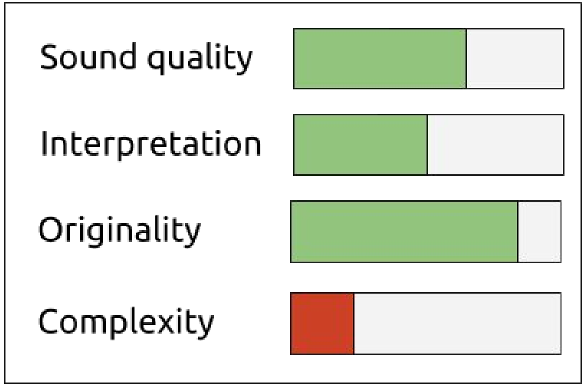
Approach useful to study evolution of music genres

Strong correlation between average sentiment and average rating

Again useful insights for musicologists

Opinionated Music Recommendation

Cluster aspects into topics and average scores for each item



Ruihai Dong, Michael P O'Mahony, and Barry Smyth (2014). Further Experiments in Opinionated Product Recommendation. In ICCBR'14, pages 110–124.

Generation of Annotated Corpus

Identify song names: Entity Linking

Identify aspects and opinion words: Aspect-based Sentiment Analysis

Associate aspects to songs

- E.g. Identify songs associated to different arousal parameters, instrument properties

Create annotations useful for MIR classifiers (e.g. mood, instruments)

Tools

Alchemy API
http://www.alchemyapi.com/products/alchemylanguage/entity-extraction

AYLIEN API http://aylien.com/text-api

Stanford NLP http://nlp.stanford.edu:8080/sentiment/rntnDemo.html

Gensim python library https://radimrehurek.com/gensim/

Senti WordNet http://sentiwordnet.isti.cnr.it/

References

Pang, B., & Lee, L. (2006). *Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, 1(2). 91–231.*

Tata, S., & Di Eugenio, B. (2010). Generating Fine-Grained Reviews of Songs from Album Reviews. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, (July), 1376–1385.

Ruihai Dong, Michael P O'Mahony, and Barry Smyth (2014). Further Experiments in Opinionated Product Recommendation. In ICCBR'14, pages 110–124.

Oramas S., Espinosa-Anke L., Lawlor A., Serra X., Saggion H. (2016). *Exploring Music Reviews for Music Genre Classification and Evolutionary Studies*. 17th International Society for Music Information Retrieval Conference. ISMIR 2016.

Dominique Moisi. The Geopolitics of Emotion: How Cultures of Fear, Humiliation, and Hope are Reshaping the World. Anchor Books, New York, NY, USA, 2010.

Outline

- Introduction to NLP
- Information Extraction
    - Construction of Music Knowledge Bases
    - Applications in MIR
- Topic Modeling
- Sentiment Analysis
- **Lexical Semantics**

Lexical Semantics

Introduction

· "What is it about the representation of a lexical item that gives rise to sense extensions and to the phenomenon of logical polysemy?" - *Pustejovsky, 1995. Introduction: Lexical Semantics in Context, Journal of Semantics.*

· Lexical Semantics is about understanding the "units of meaning" of the language. Not only words, but also compound words, phrases, affixes, etc.

· In NLP, two main approaches for understanding this meaning: Pattern-based and **distributional semantics**.

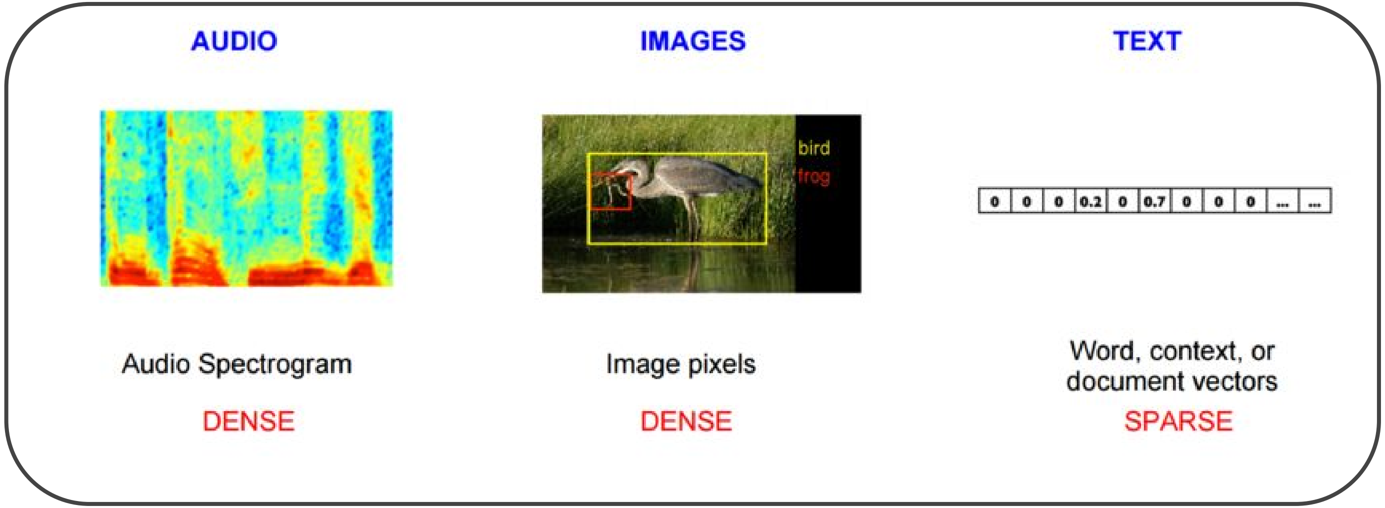
· Distributional semantics intersects with *Relational Semantics*, i.e. establishing relationships between pairs of lexical units.

Distributional Lexical Semantics

- “You shall know a word by the company it keeps”, Firth (1957).

Distributional Lexical Semantics

- “You shall know a word by the company it keeps”, Firth (1957).
- Project linguistic items in a vector space.



https://www.tensorflow.org/versions/r0.10/tutorials/word2vec/index.html

Distributional Lexical Semantics

- “You shall know a word by the company it keeps”, Firth (1957).
- Project linguistic items in vector space.
- Count-based vs Predictive models (Baroni et al., ACL 2014).

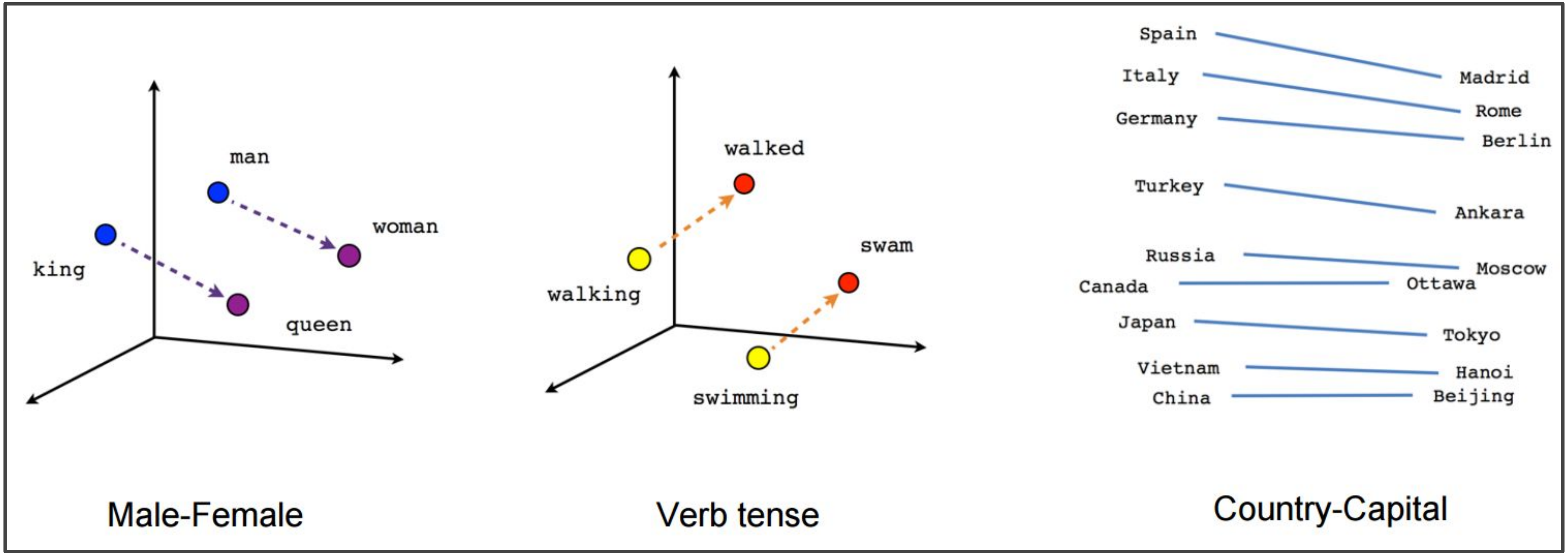
  * Count-based: `freq(x,y)` in corpus and map these stats to a dense vector.

  * Predictive: Try to predict a word from its neighbours.

  → *the quick brown fox jumped over the lazy dog*

  ⇒ ([the, brown], **quick**), ([quick, fox], **brown**), ([brown, jumped], **fox**), ...

Distributional Lexical Semantics



Distributional Lexical Semantics

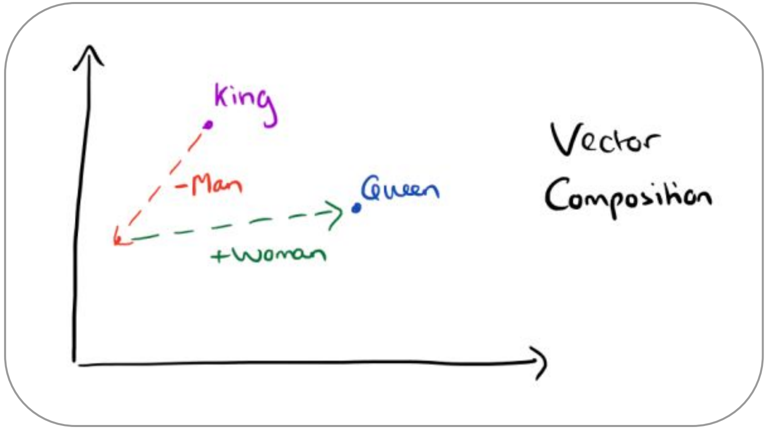
```
>>> from gensim.models import Word2Vec

>>> model = Word2Vec.load(PATH)
```

gensim

https://radimrehurek.com/gensim/models/word2vec.html

Distributional Lexical Semantics

- Famous analogy example




```
>>> model.most_similar(positive=['woman', 'king'], negative=['man'])
```

[(u'queen’, 0.71), (‘monarch’, 0.61), (u'princess', 0.59) …]

Distributional Lexical Semantics

- Can be used to discover facts about music. **Representative instruments!**

*Hendrix is to guitar as Mozart is to **x***

Distributional Lexical Semantics

- Can be used to discover facts about music. **Representative instruments**

Hendrix is to guitar as Mozart is to ***x***

```
>>> model.most_similar(positive=['Mozart', 'guitar'], negative=['Hendrix'])
```

[(u'piano', 0.52), (u'accordion', 0.47), (u'mandolin', 0.47), (u'banjo', 0.47), (u'trombone', 0.46), (u'flute', 0.44) …]

Distributional Lexical Semantics

- Can be used to discover facts about music. **Associated Music Genres**

Enrique Iglesias is to Pop as Elvis Presley is to …

```
model.most_similar(positive=['Elvis', 'Pop'], negative=['Enrique_Iglesias'])
```

Distributional Lexical Semantics

- Can be used to discover facts about music. **Associated Music Genres**

Enrique Iglesias is to Pop as Elvis Presley is to …

```
model.most_similar(positive=['Elvis', 'Pop'], negative=['Enrique_Iglesias'])
```

```
[(u'Country', 0.57), (u'Rock', 0.57), (u'Reggae', 0.57), (u'Blues', 0.55), (u'Metal', 0.55), (u'Jazz', 0.54), (u'Punk', 0.54), (u'Hip_Hop', 0.54), (u'Rap', 0.53), (u'Bluegrass', 0.53)]
```

Innovative Distributional Representations

· Lots of work in enhancing vector representations of (not only) language.

· **Paragraph Vector** (Le and Mikolov, 2014 ICML)

· BabelNet ***synsets*** and **Wikipedia pages** (Camacho-Collados, 2015 NAACL; Iacobacci et al. 2015 ACL)

· **Twitter emojis** (Barbieri et al., 2016 LREC)

· **Retrofitting** word vectors to semantic lexicons (Faruqui et al., 2015 NAACL)

* Forcing synonymy, hypernymy, meronymy, or collocational information.

· Domain-specific models. A word2vec model in the **music domain**.

A word2vec model in the Music domain

· The model has a restricted vocabulary of 21635 words.

· Trained over 19850433 raw words and 861414 sentences.

· Trained on the following datasets (overall +72k documents):

* Grove music encyclopedia, 16708 biographies.

* Last.fm, 23015 biographies.

* Songfacts trivia, biographies and tidbits, 32326 documents.

* Available at (we will upload further versions trained on larger corpora and additional preprocessing): http://mtg.upf.edu/nlp-tutorial

A word2vec model trained on music corpora

>>> model.most_similar(positive=["**beatles**","**mick_jagger**"],negative=["**john_lennon**"])

[(**u'rolling_stones'**, 0.6256111860275269), …]

>>> model.most_similar(positive=["**dance-pop**","**zz_top**"],negative=["**lady_gaga**"])

[(**u'jazz-rock'**, 0.6238052845001221) …]

>>> model.most_similar(positive=["**syd_barrett**","**roger_waters**"])

[(**u'david_gilmour'**, 0.7655651569366455) …]

>>> model.most_similar(positive=["**iggy_pop**"])

[(**u'patti_smith'**, 0.7802923917770386) …]

References

WordNet: Miller, G. A. (1995). WordNet: a lexical database for English. Communications of the ACM, 38(11), 39-41.

Firth's paper: Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955.

Count-based vs Predictive: Baroni, M., Dinu, G., & Kruszewski, G. (2014, June). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In ACL (1) (pp. 238-247).

Word2Vec: Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems (pp. 3111-3119).

SensEmbed: Iacobacci, I., Pilehvar, M. T., & Navigli, R. (2015). SensEmbed: learning sense embeddings for word and relational similarity. In Proceedings of ACL (pp. 95-105).

SensEmbed for Disambiguation: Bovi, C. D., Anke, L. E., & Navigli, R. (2015). Knowledge Base Unification via Sense Embeddings and Disambiguation. In Proceedings of EMNLP (pp. 726-736).

SensEmbed for Taxonomy Learning: Espinosa-Anke, L., Saggion, H., Ronzano, F., & Navigli, R. (2016). ExTaSem! Extending, Taxonomizing and Semantifying Domain Terminologies. AAAI 2016.

SensEmbed for Artist Similarity: Oramas, S., Sordo, M., Espinosa-Anke, L., & Serra, X. (2015). A Semantic-based Approach for Artist Similarity. ISMIR 2015.

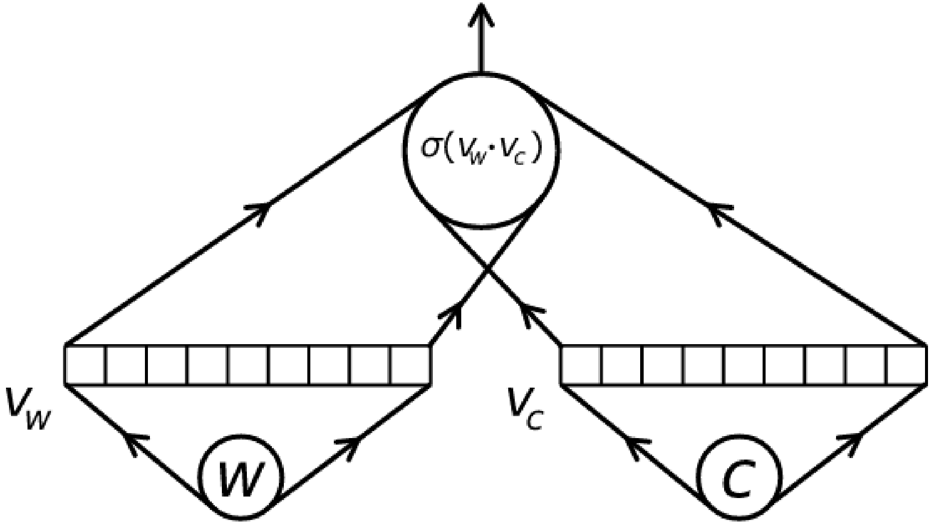
Other Sense-level Vectors: Camacho-Collados, J., Pilehvar, M. T., & Navigli, R. (2015). NASARI: a novel approach to a semantically-aware representation of items. In Proceedings of NAACL (pp. 567-577).

Word2vec and beyond

Word2vec

Predict a context word **c** ∈ ($\mathbf{w}_{i-L}$, . . . , $\mathbf{w}_{i-1}$, $\mathbf{w}_{i+1}$, . . . , $\mathbf{w}_{i+L}$) given a word $\mathbf{w}_i$

$$P(D = 1|w, c) = \sigma(v_w \cdot v_c)$$



$$\sigma(a) = \frac{1}{1+e^{-a}}$$

Skip-Gram Negative Sampling (SGNS)

Maximize $P(D = 1|w, c)$ for observed (w, c)

Maximize $P(D = 0|w, c)$ for randomly sampled “negative” examples (w , c)

$$\arg\max_\theta \sum_{(w,c)\in D} \log \sigma(v_c \cdot v_w) + \sum_{(w,c)\in D'} \log \sigma(-v_c \cdot v_w)$$

Word2vec as Matrix Factorization

Word and context embeddings matrices **W** and **C** are learnt

W is typically used in NLP, while **C** is ignored

$\mathbf{C} \cdot \mathbf{W}^{\mathrm{T}} = \mathbf{M}$ what is **M**?

According to Levy et al. 2014

$$M_{ij}^{\text{SGNS}} = W_i \cdot C_j = \vec{w}_i \cdot \vec{c}_j = PMI(w_i, c_j) - \log k$$

$$PMI(x, y) = \log \frac{P(x, y)}{P(x)P(y)}$$

Factoring PMI with SVD

$$PMI(w,c) = \log \frac{\#(w,c) \cdot |D|}{\#(w) \cdot \#(c)}$$

Factoring PMI with SVD

$$PMI(w,c) = \log \frac{\#(w,c) \cdot |D|}{\#(w) \cdot \#(c)}$$

$$PPMI(w,c) = \max\left(PMI\left(w,c\right),0\right)$$

Factoring PMI with SVD

$$PMI(w,c) = \log \frac{\#(w,c) \cdot |D|}{\#(w) \cdot \#(c)}$$

$$PPMI(w,c) = \max\left(PMI\left(w,c\right),0\right)$$

$$M_d = U_d \cdot \Sigma_d \cdot V_d^\top$$

d dimensions

Factoring PMI with SVD

$$PMI(w,c) = \log \frac{\#(w,c) \cdot |D|}{\#(w) \cdot \#(c)}$$

$$PPMI(w,c) = \max\left(PMI\left(w,c\right),0\right)$$

$$M_d = U_d \cdot \Sigma_d \cdot V_d^\top$$

d dimensions

$$W^{\text{SVD}_{1/2}} = U_d \cdot \sqrt{\Sigma_d} \qquad C^{\text{SVD}_{1/2}} = V_d \cdot \sqrt{\Sigma_d}$$

SGNS vs SVD-PMI

- Factorized PMI is not exactly the same as Word2vec with SGNS
- Factorized PMI is easier and faster
- Both have similar performance in similarity tasks
- Word2vec with SGNS better performance in some NLP tasks (analogies)

More about Word2vec

https://www.tensorflow.org/versions/r0.10/tutorials/word2vec/index.html

http://cgi.cs.mcgill.ca/~enewel3/posts/implementing-word2vec/

http://hduongtrong.github.io/2015/11/20/word2vec/

T. Mikolov et al (2013): Distributed Representations of Words and Phrases and their Compositionality. Advances in neural information processing systems.

O. Levy, Y. Goldberg (2014): Neural Word Embedding as Implicit Matrix Factorization. NIPS 2014

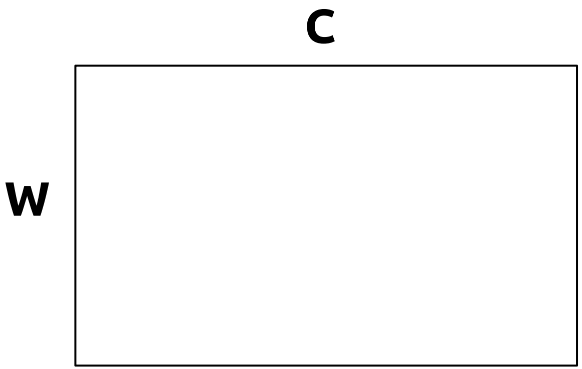
Beyond words

C and **W** can be different from words

Ej.:

W songs or artists, **C** playlists

W tags, **C** items



We can learn vector embeddings of musical items

Word2vec in Playlists

Trained with Gensim in Art of the Mix playlists
(http://labrosa.ee.columbia.edu/projects/musicsim/aotm.htm)

model.most_similar('miles davis')
[('john clotrane', 0.88384414), ('dizzie gillespie', 0.78484219), ('charlie walker', 0.74520659)]

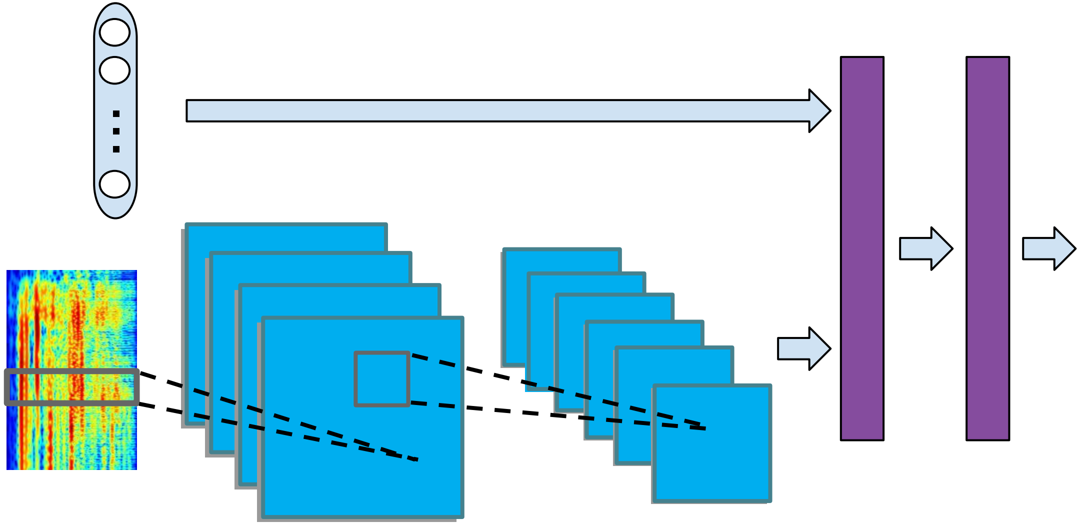
model.most_similar('marilyn manson')
[('godsmack', 0.93274206), ('white zombie', 0.91064525), ('drowning pool', 0.90275443)]

model.most_similar('nirvana')
[('soundgarden', 0.84231329), ('pearl jame', 0.8271907), ('oysterhead', 0.81855756)]

Word2vec for Deep Learning

Word embeddings are ideal input for deep neural networks

Audio and word embeddings can be combined in a deep neural network



Conclusions

- The extraction of high level semantic representations from text have been shown useful in different MIR problems.
- There is a need of development of new methodologies that exploit these semantic representations in MIR.
- Word Embeddings and Deep Learning opens a new world of already unexploited possibilities for multimodal approaches.
- This tutorial has been an initial attempt to boost the interaction between the NLP and MIR communities. Future objectives:
    - Creation of challenges
    - Attract the NLP community to MIR problems and datasets

Thanks!

Questions? Ideas? Suggestions?

@sergiooramas @luisanke @zangsir