Natural Language Processing for Music Information Retrieval

Sergio Oramas, Luis Espinosa-Anke
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

Horacio Saggion
TALN, Universitat Pompeu Fabra
Barcelona, Spain

Xavier Serra
MTG, Universitat Pompeu Fabra
Barcelona, Spain

EXCELENCIA
MARÍA
DE MAEZTU

Music meets NLP

upf. Music Technology Group

taln research group

Objectives

Provide a general introduction to NLP.

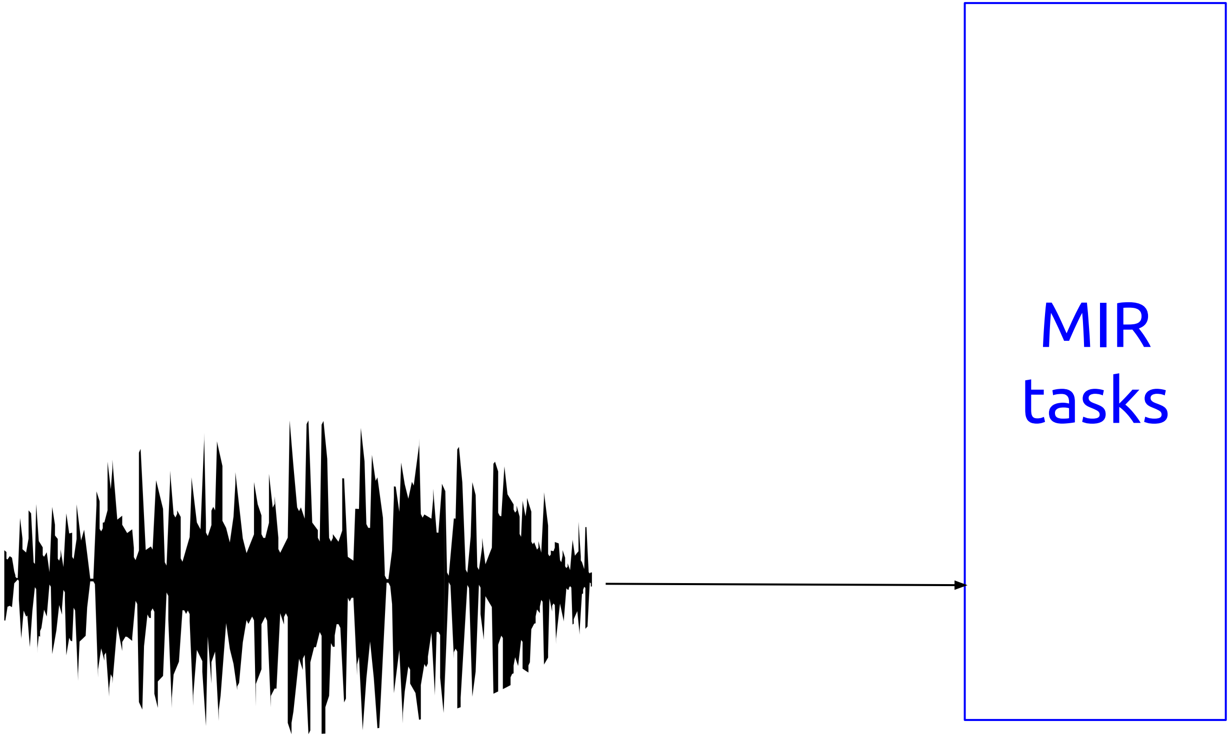
Identify areas of NLP with potential application in MIR.

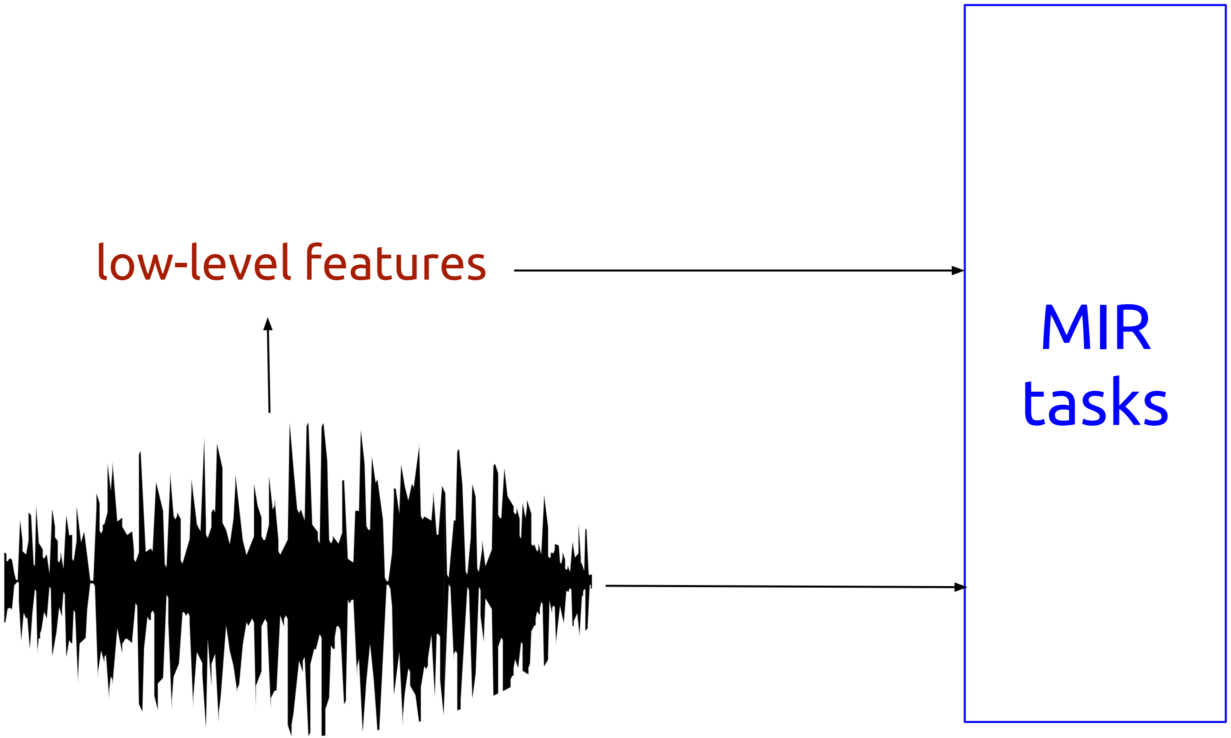
Address the extraction of semantic information from music text corpora.

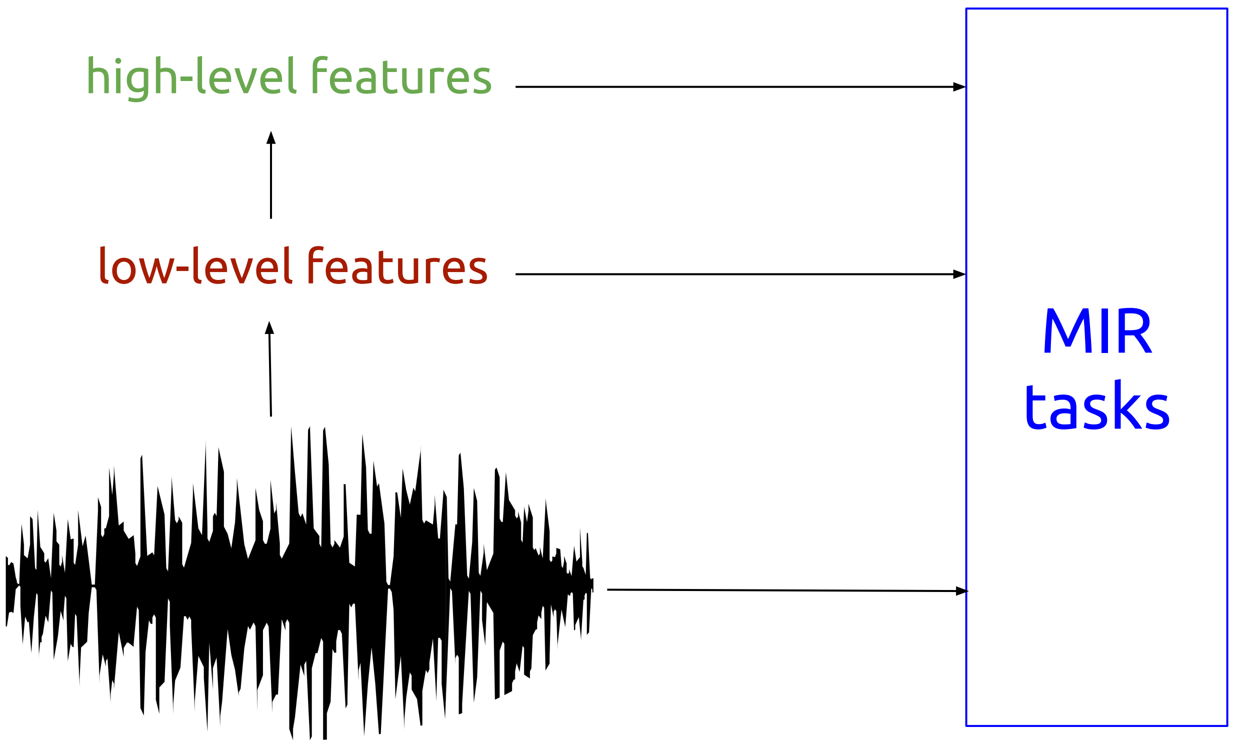
Show methodologies for exploiting semantic information in MIR.

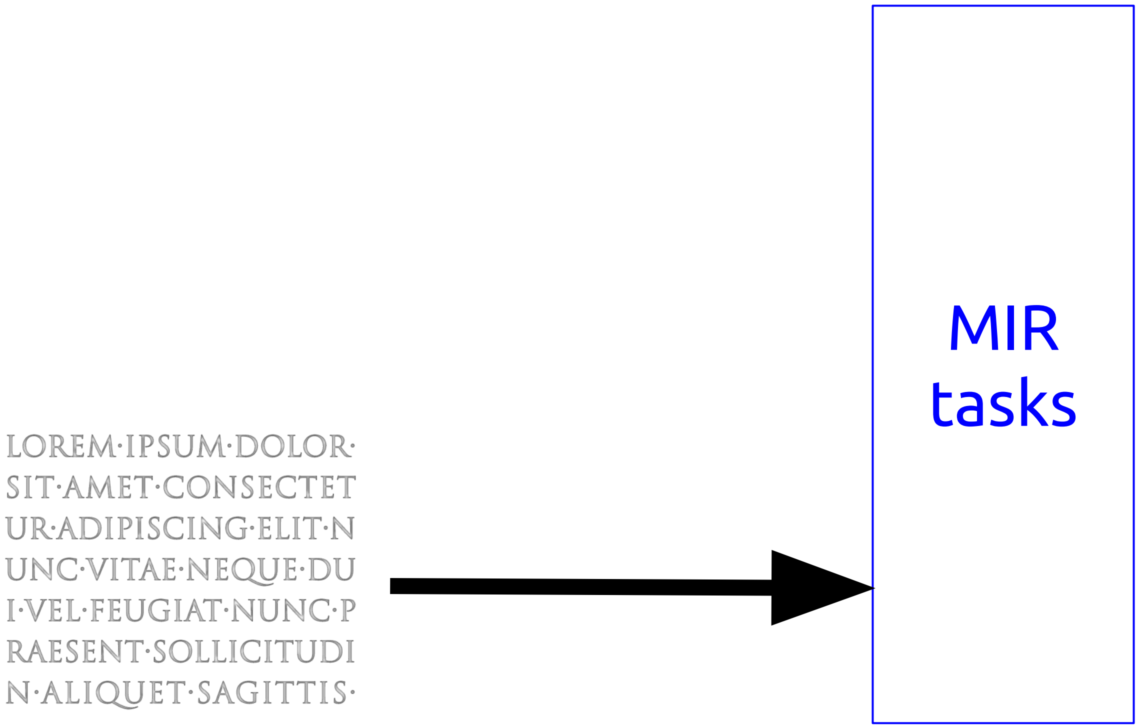
Illustrate latest tendencies in NLP

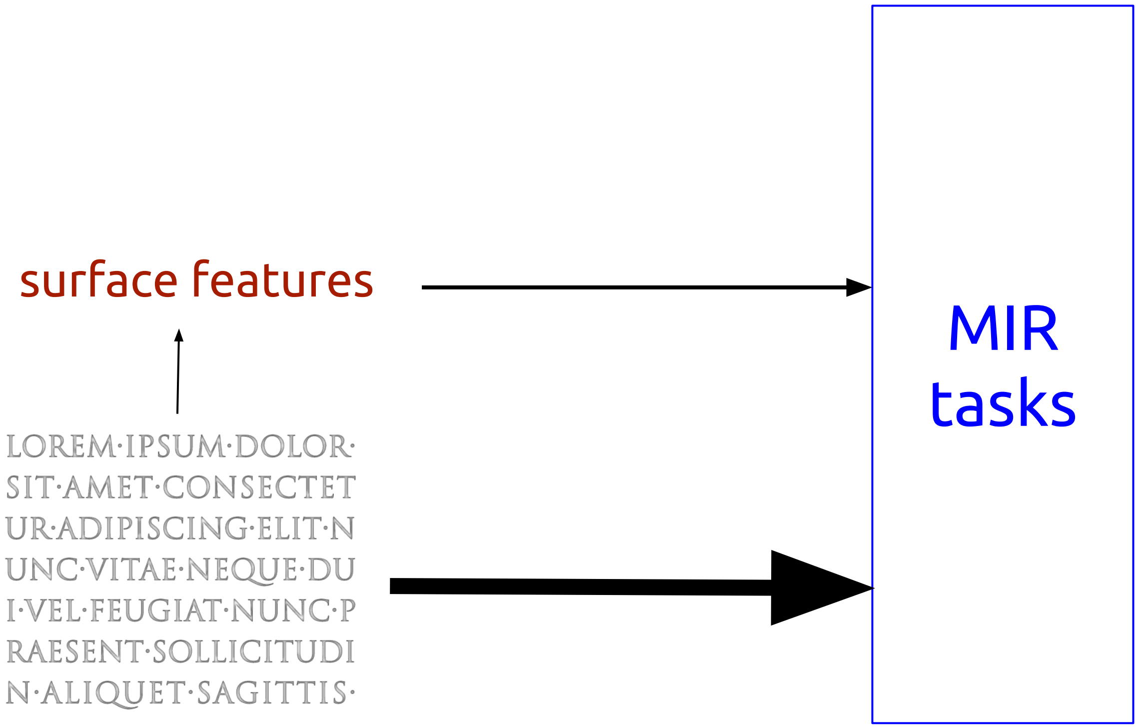
Why semantic information?


MIR
tasks

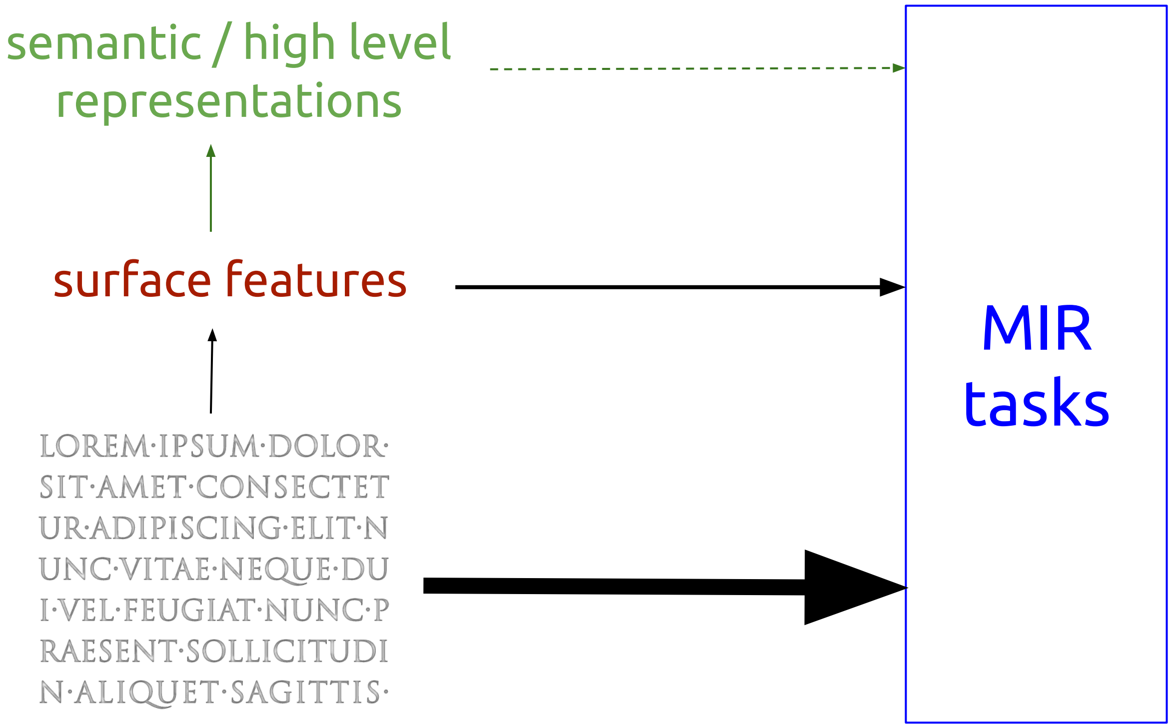

low-level features
MIR tasks


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

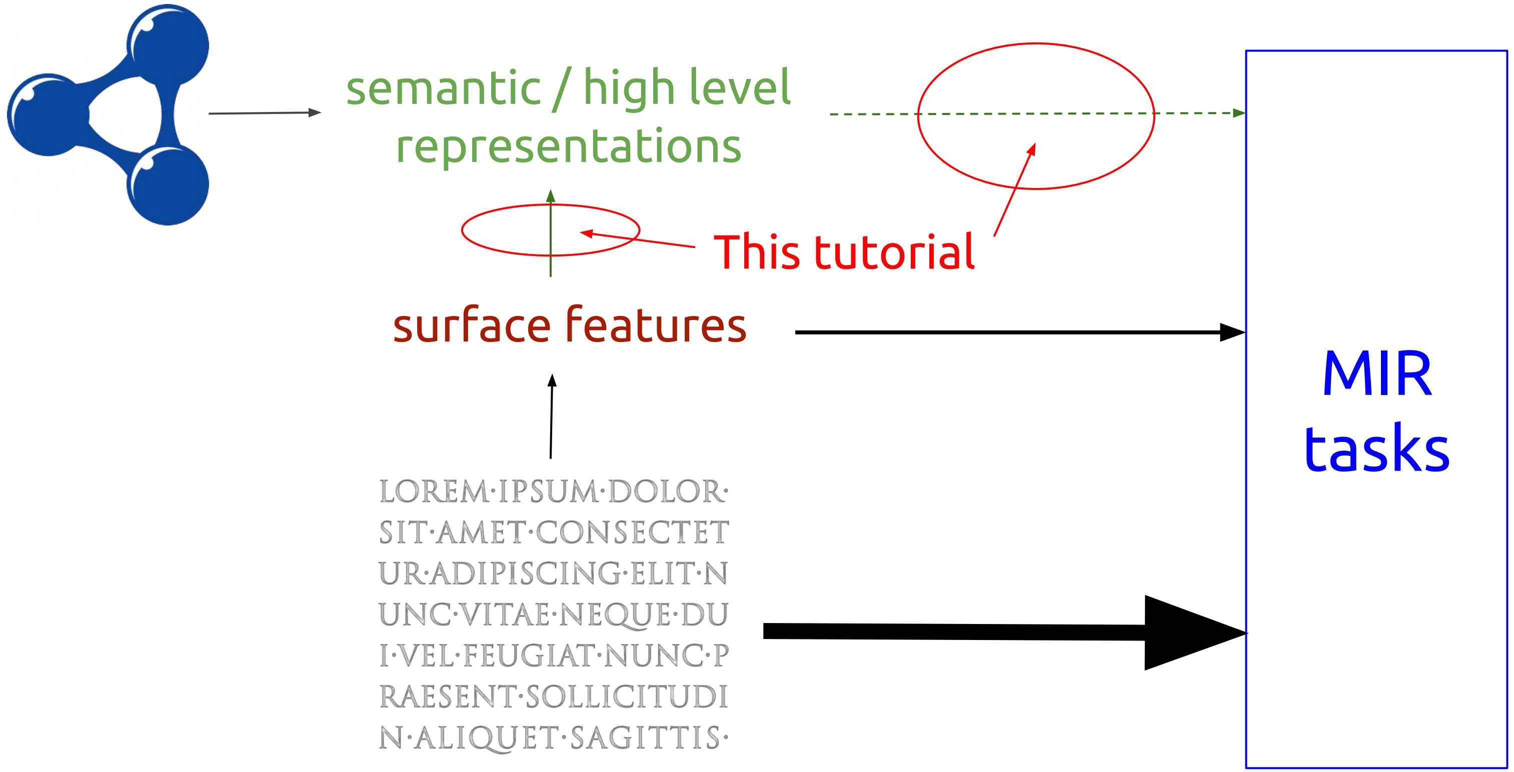


Peter Knees & Markus Schedl (2013): A Survey of Music Similarity and Recommendation from Music Context Data. ACM-TOMM.


semantic / high level representations
surface features
LOREM·IPSUM·DOLOR·SIT·AMET·CONSECTETUR·ADIPISCING·ELIT·NUNC·VITAE·NEQUE·DUI·VEL·FEUGIAT·NUNC·PRAESENT·SOLLICITUDIN·ALIQUET·SAGITTIS·
MIR tasks

semantic / high level
representations

MIR
tasks


semantic / high level 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

Corpora in MIR Related Work

Lyrics

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

Corpora in MIR Related Work

Lyrics

This tutorial

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

Outline

- Introduction to NLP (20 mins)
- Information Extraction (10 mins)
    - Construction of Music Knowledge Bases (15 mins)
    - Semantic Enrichment of Musical Texts (5 mins)
- Applications in MIR (25 mins)

--- break ---

- Applications in Musicology (10 mins)
- Lexical Semantics (15 mins)
- Deep Learning (10 mins)
- Conclusions and Future (5 mins)

Outline

- **Introduction to NLP**
- Information Extraction
  - Construction of Music Knowledge Bases
  - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- Conclusions and Future

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.

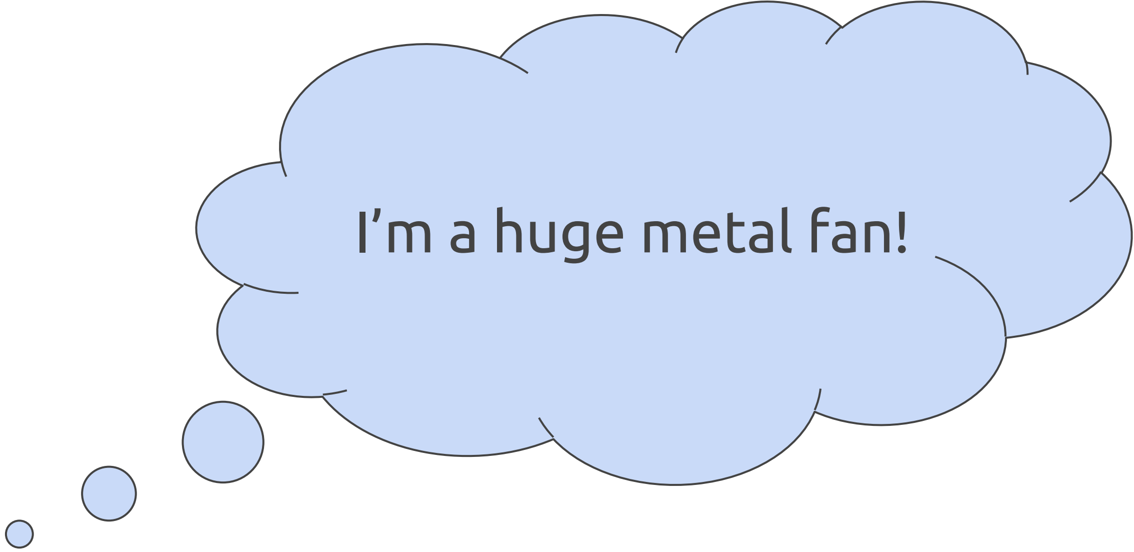
What is Natural Language Processing?

- 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!

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)

Successful NLP: "Will a computer program ever be able to convert a piece of English text into a programmer friendly data structure that describes the meaning of the natural language text? Unfortunately, **no consensus has emerged about the form or the existence of such a data structure** "(Collobert et al., 2011).

Core elements in NLP - Part-of-Speech Tagging

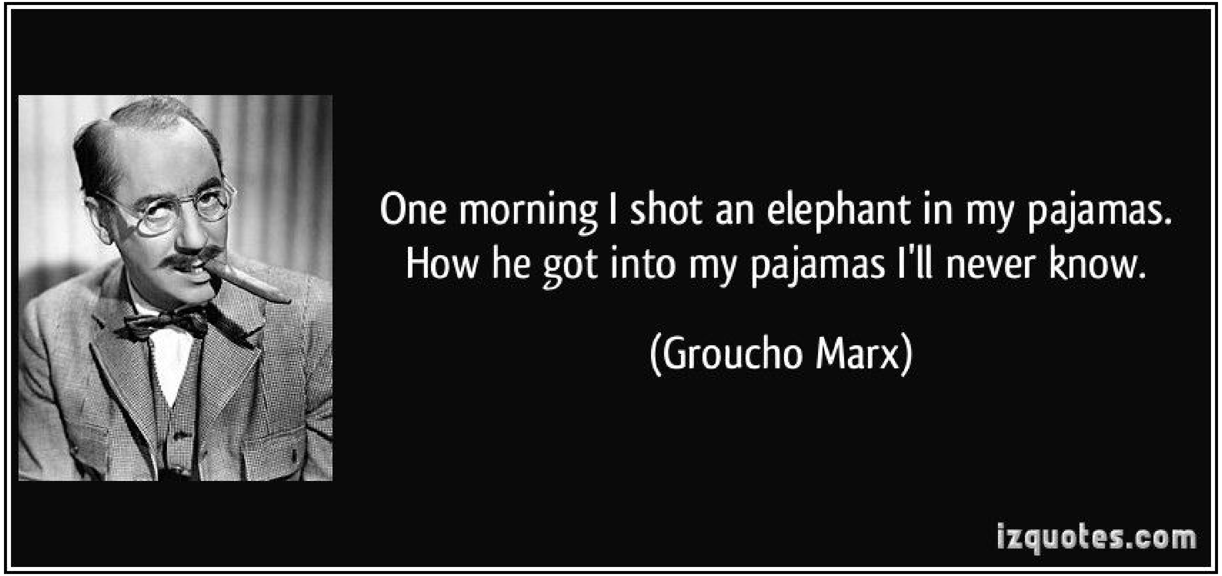
I like jazz music, it’s like being alive for a second.

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*.

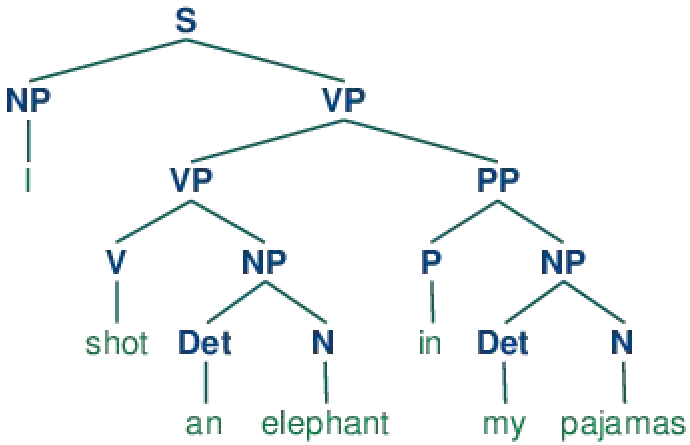
- By analyzing sentence structure, we understand the underlying meaning in a sentence.

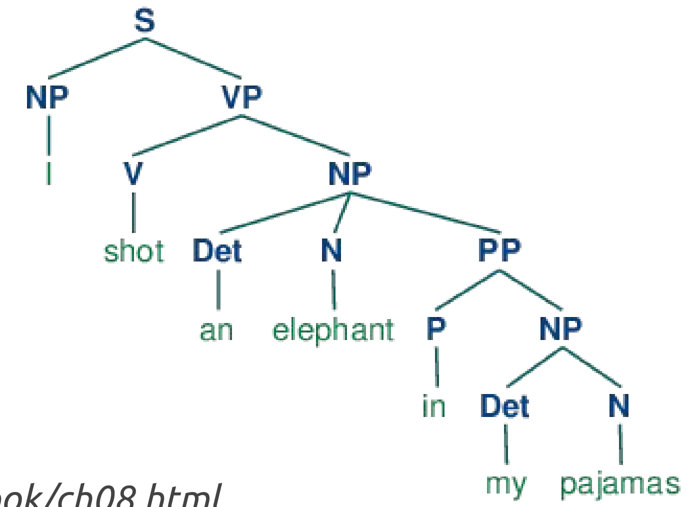
http://www.nltk.org/book/ch08.html

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.



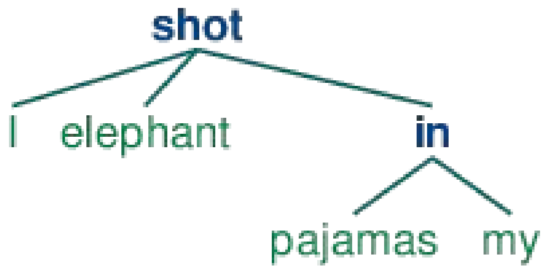


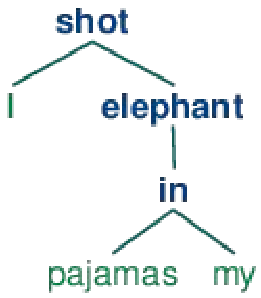
http://www.nltk.org/book/ch08.html

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 keyboard player born in South Africa 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.

http://nlp.stanford.edu/projects/coref.shtml

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

  * Author Profiling

  * Machine Translation

  * Sentiment Analysis

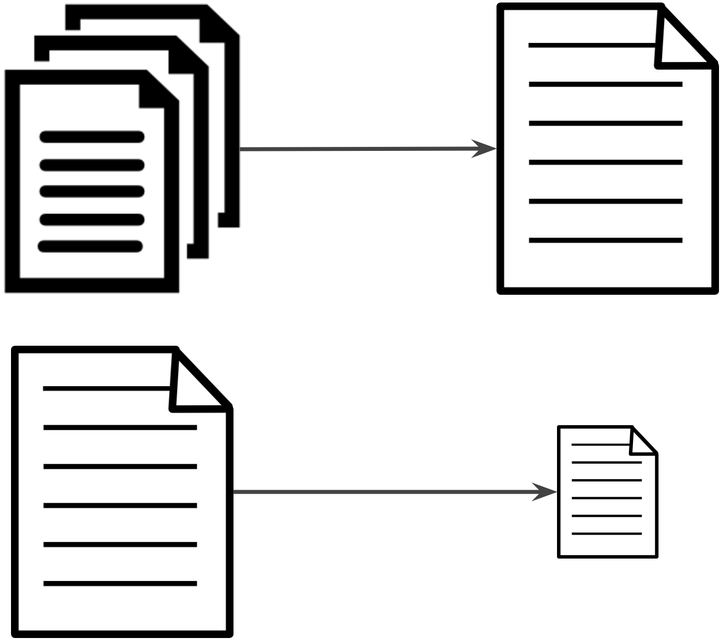
NLP Tasks - Summarization

- 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*.

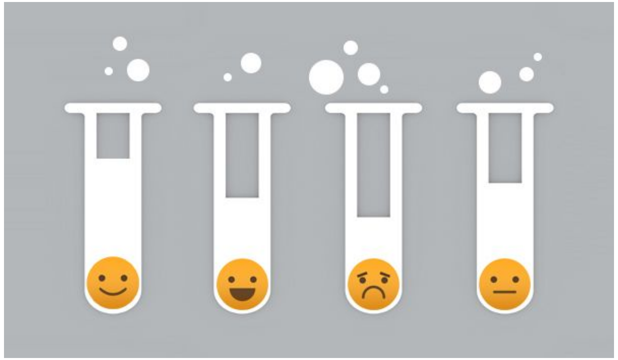
Sentiment Analysis

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



Sentiment Analysis

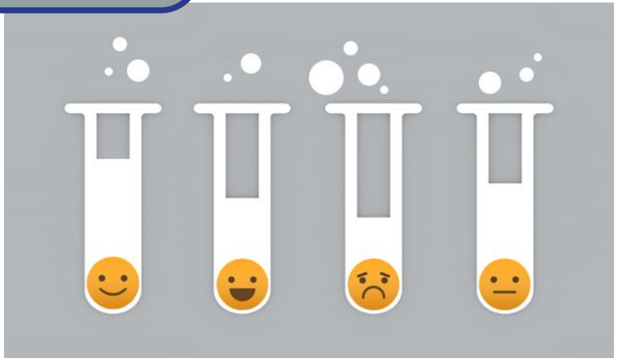
Computational study of opinions, sentiments, subjectivity, evaluations, attitudes, appraisal, affects, views, emotions, etc., expressed in text.

Complex NLP task

go read the book!

Pang, B., & Lee, L. (2006). *O* *formation Retrieval, 1(2). 91–231.*

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



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**

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/

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/

· **Deep Learning:**

- **Keras https://keras.io/**
- **Tflearn http://tflearn.org/**
- **Tensorflow https://www.tensorflow.org/**
- **Theano http://deeplearning.net/software/theano/**
- **DyNet (formerly cnn) https://github.com/clab/dynet**

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

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
    - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- Conclusions and Future

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 Identification

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

Information Extraction

Entity Recognition

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

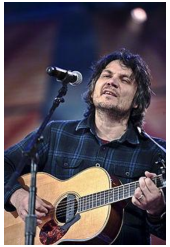
https://en.wikipedia.org/wiki/Wilco

WIKIPEDIA
The Free Encyclopedia

Entity Linking or Disambiguation

Organization

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



Work of art

Person

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

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
- Reference disambiguation

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)
  - e.g. Supertrópica is not in Wikipedia

Entity Linking

Entity linking needs to handle:

- Name variations (entities are referre
    - e.g. Elvis, Elvis Presley, Elvis Aaron Presle
- Entity ambiguity (the same string ca
    - e.g. Prince, Debut, Bach, Strauss
- Missing entities (there is no target e
    - e.g. Supertrópica is not in Wikipedia

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/

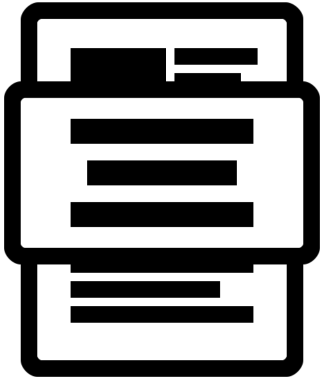
DBpedia Spotlight. Installable web service. KB: DBpedia. https://github.com/dbpedia-spotlight/dbpedia-spotlight

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

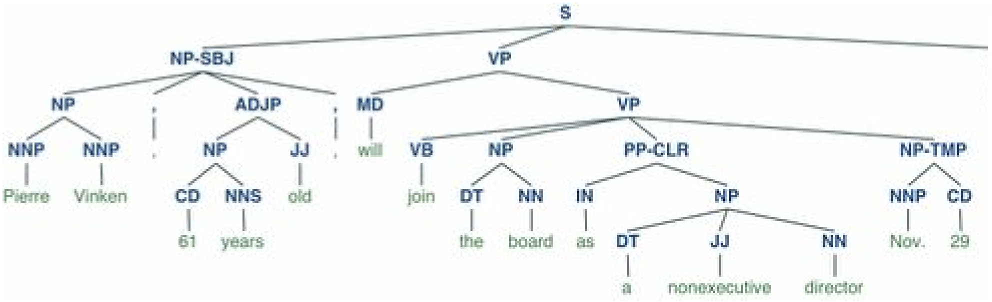


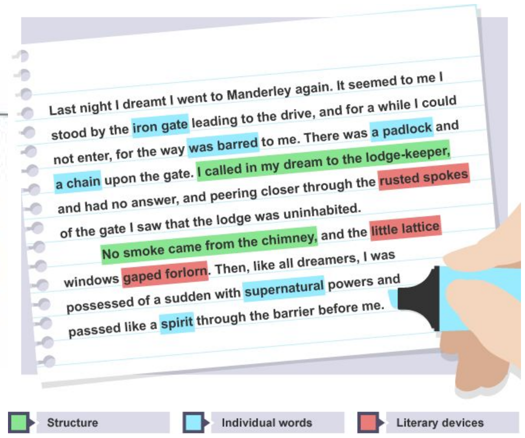
Created by Alex Getty
from Noun Project

Relation Extraction

Typical **features**:

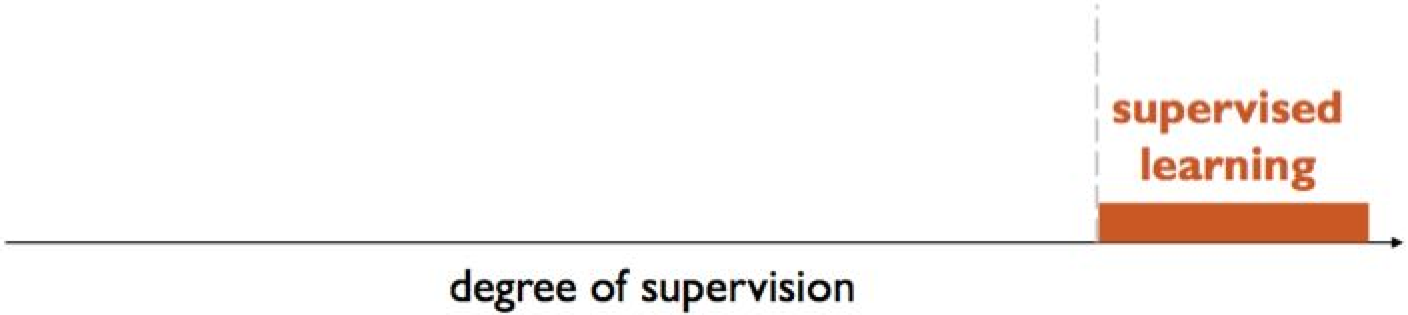
- morphologic, syntactic, semantic, statistical
- context words + part-of-speech tags, dependency paths, named entities





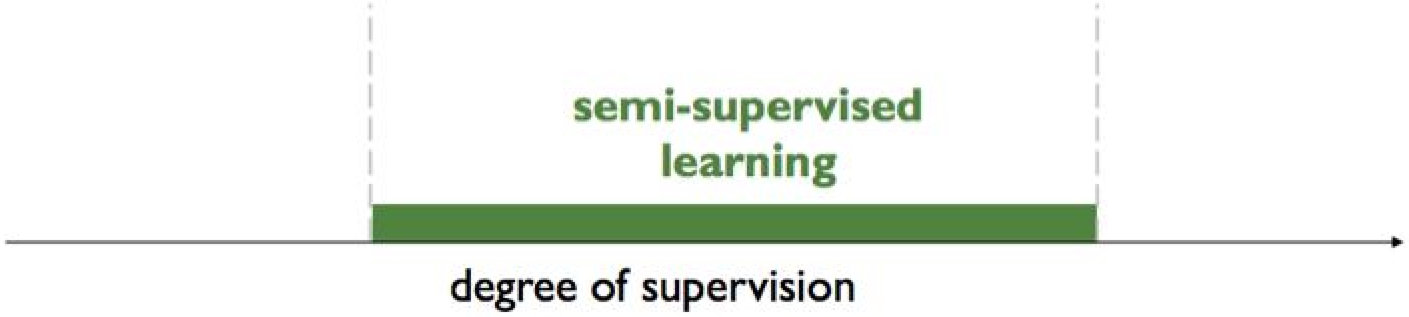
Relation Extraction

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



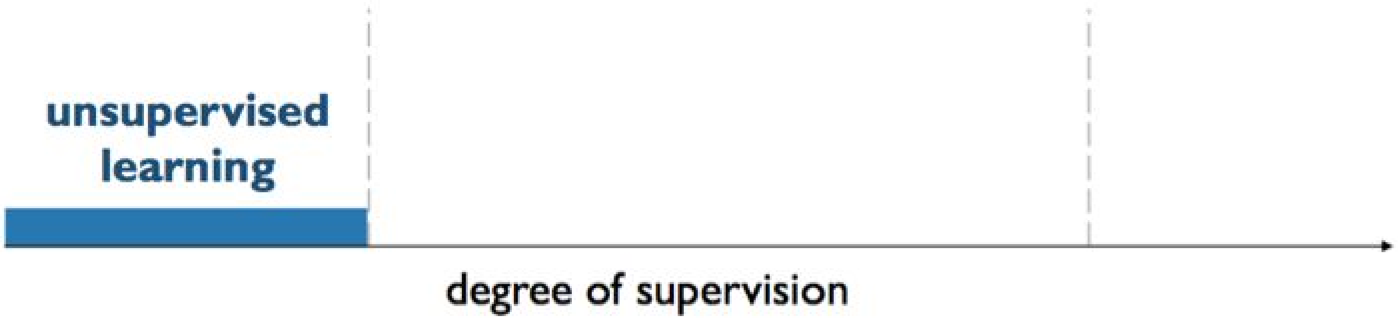
Relation Extraction

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

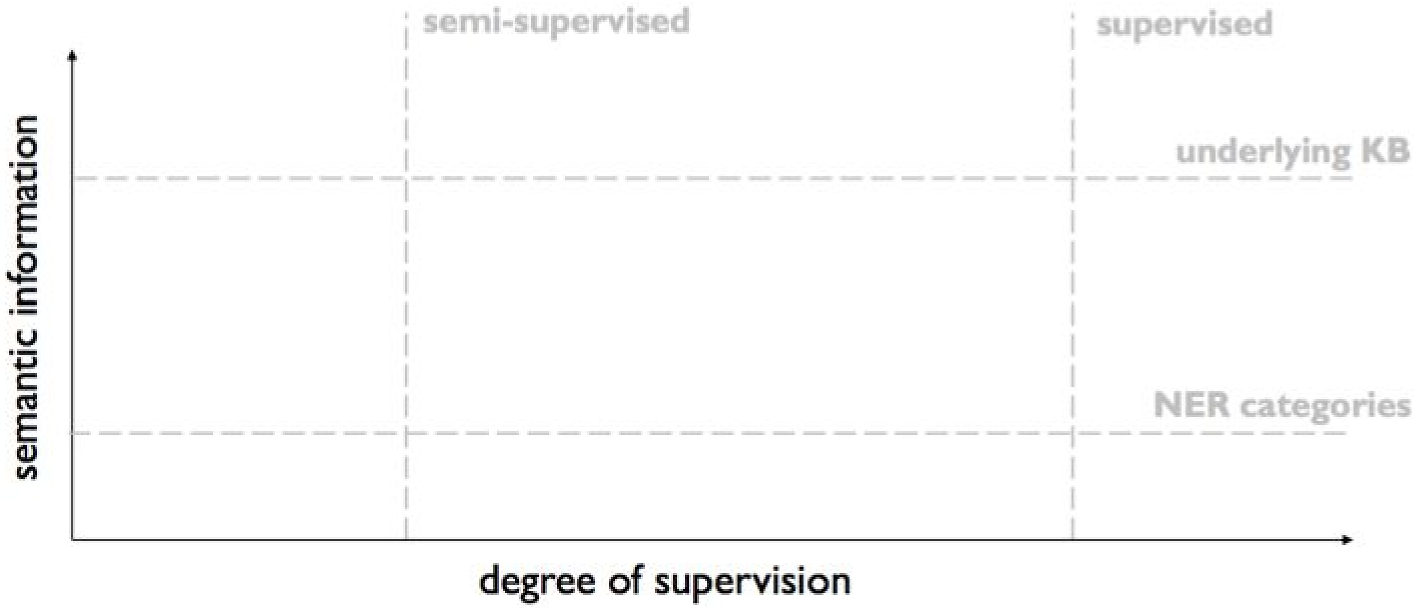


Relation Extraction

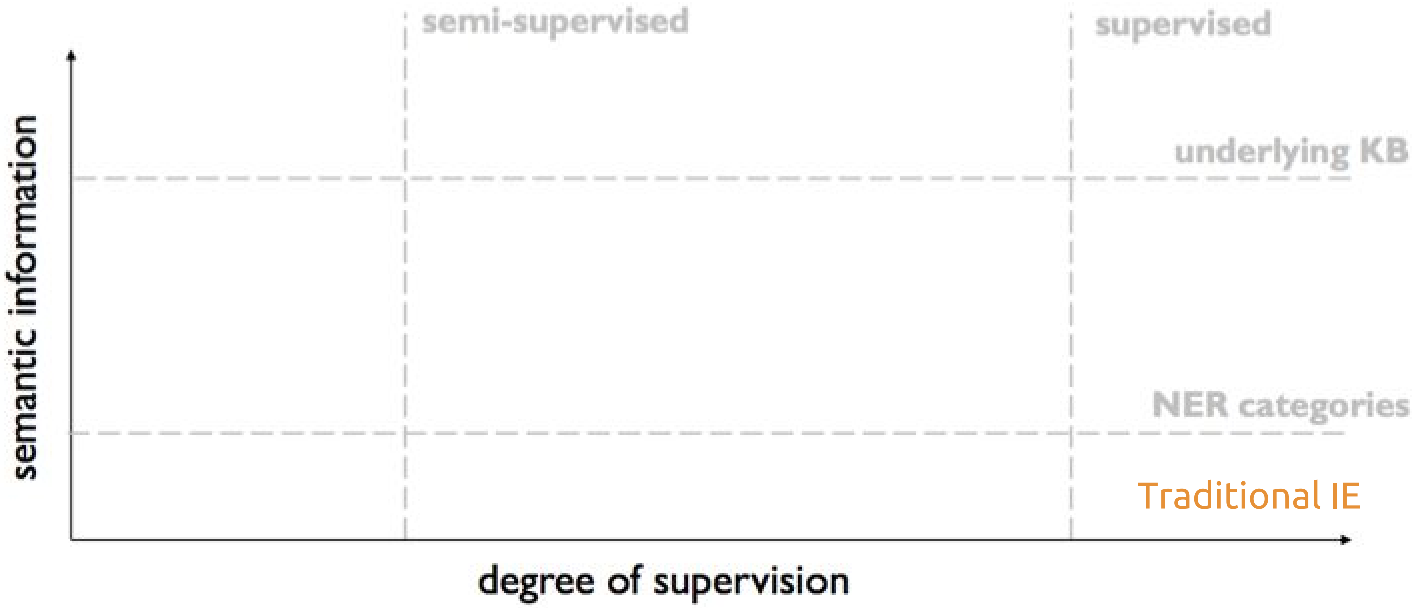
- Input:
  - Large corpus of unstructured text
  - ~~Set of semantic relations~~
  - ~~Labelled training data~~
- Output:
  - Knowledge Base of triples
  - 〈 entity, relation, entity 〉
  - Set of semantic relations



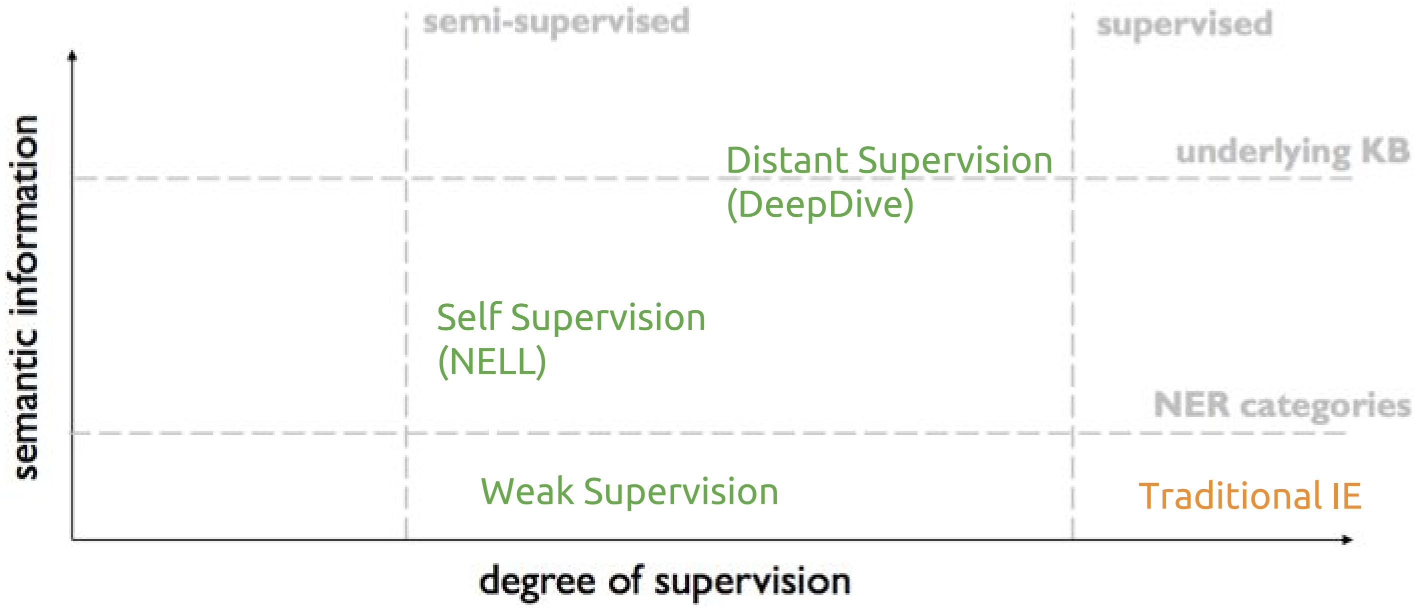
Relation Extraction



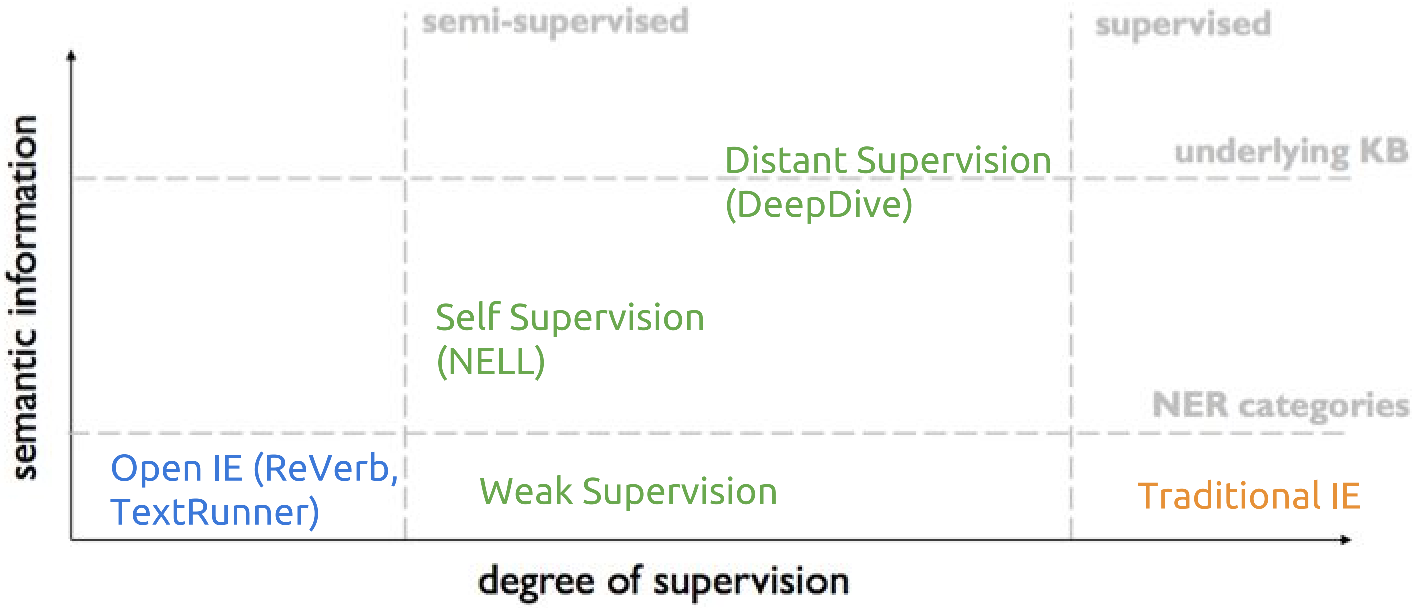
Relation Extraction



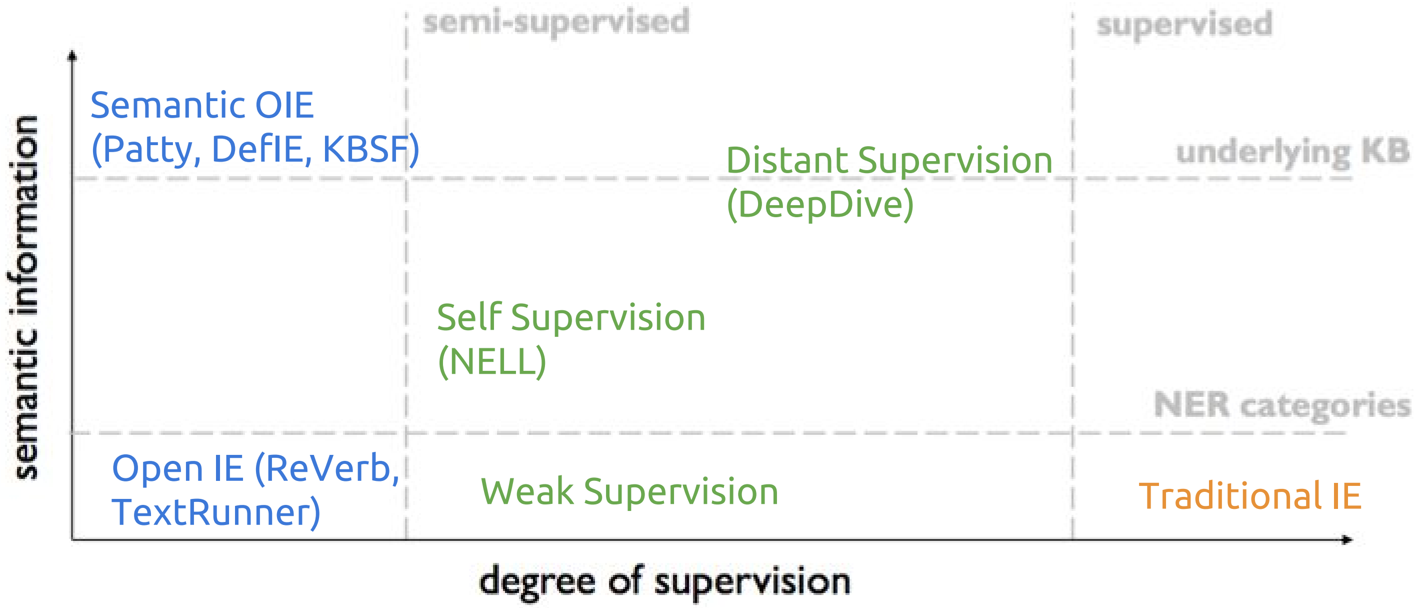
Relation Extraction



Relation Extraction

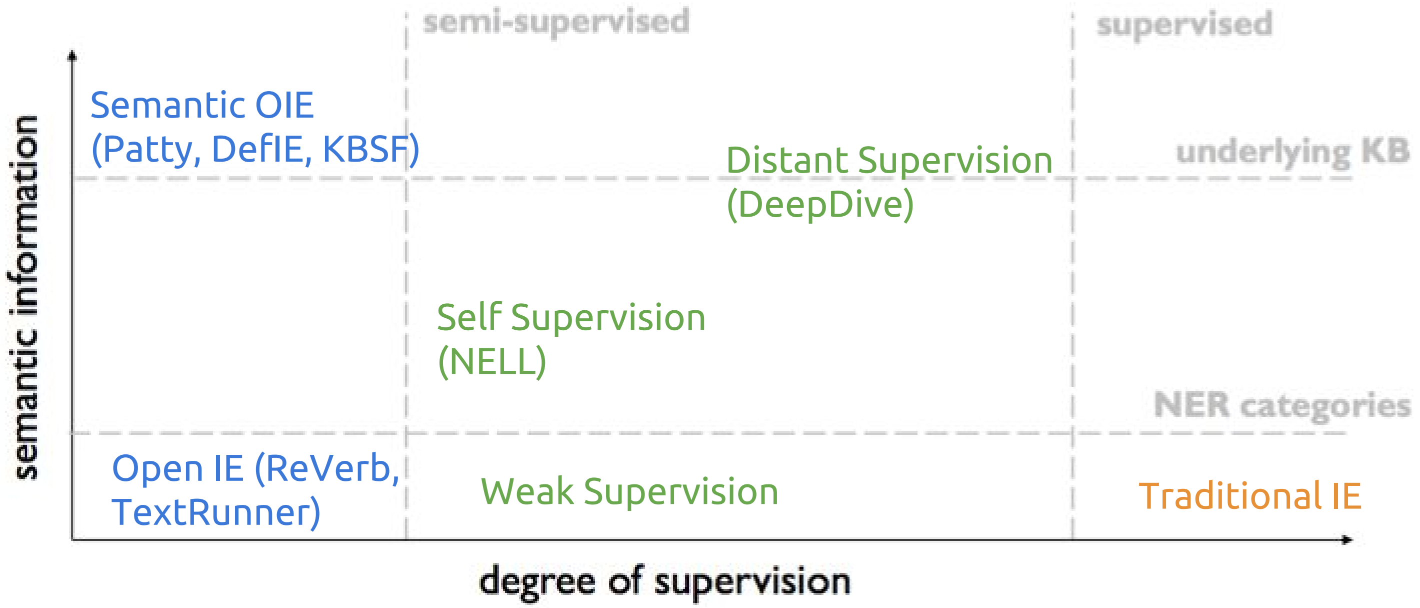


Relation Extraction



Relation Extraction

Further information in http://wwwusers.di.uniroma1.it/~dellibovi/talks/talk_OIE.pdf



Semantic Open IE

Entity Linking + Open Information Extraction

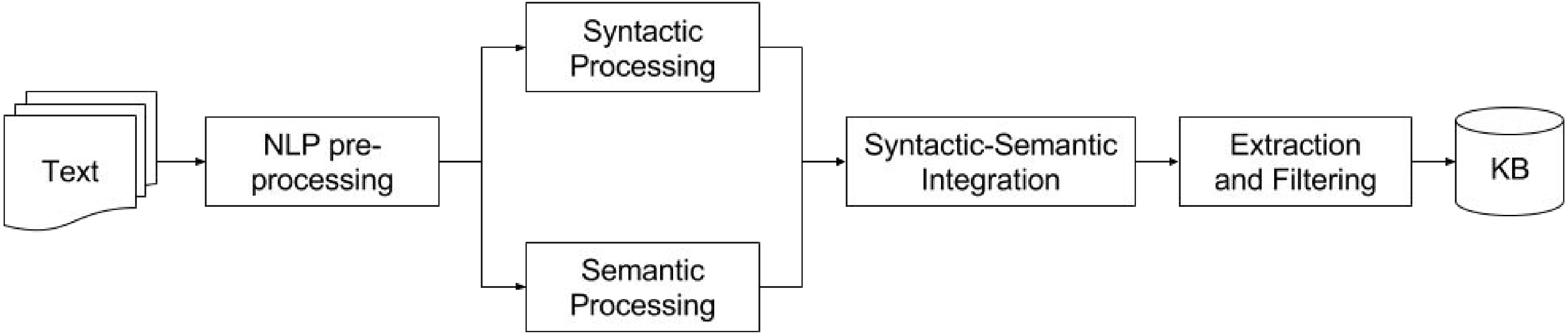
Advantages

- 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.

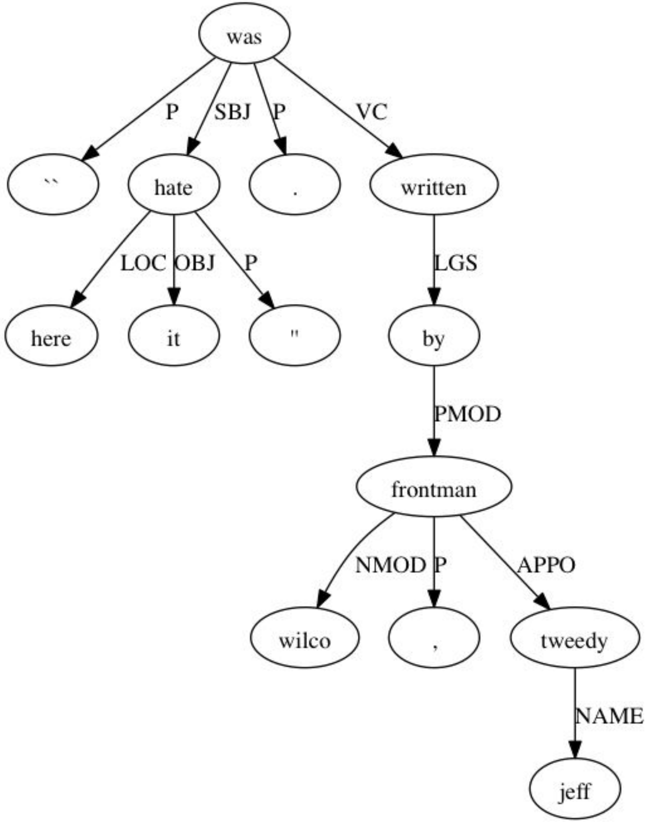
Semantic Open IE

- Entity linking -> Semantic Information
- Dependency parsing -> Syntactic Information
- Semantic-Syntactic integration
- Shortest path between entities
- Filtering of relations



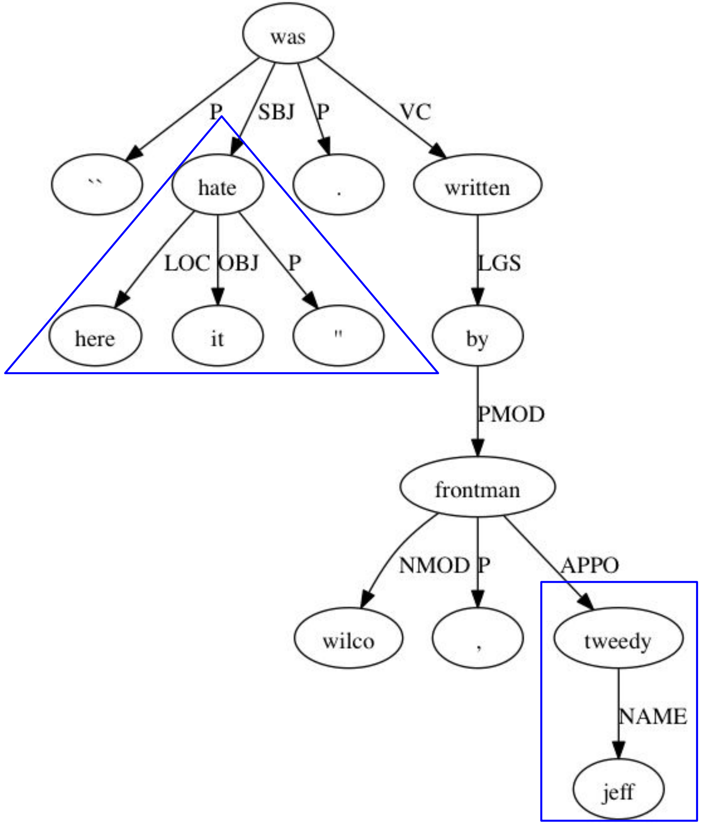
`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .

Semantic Open IE



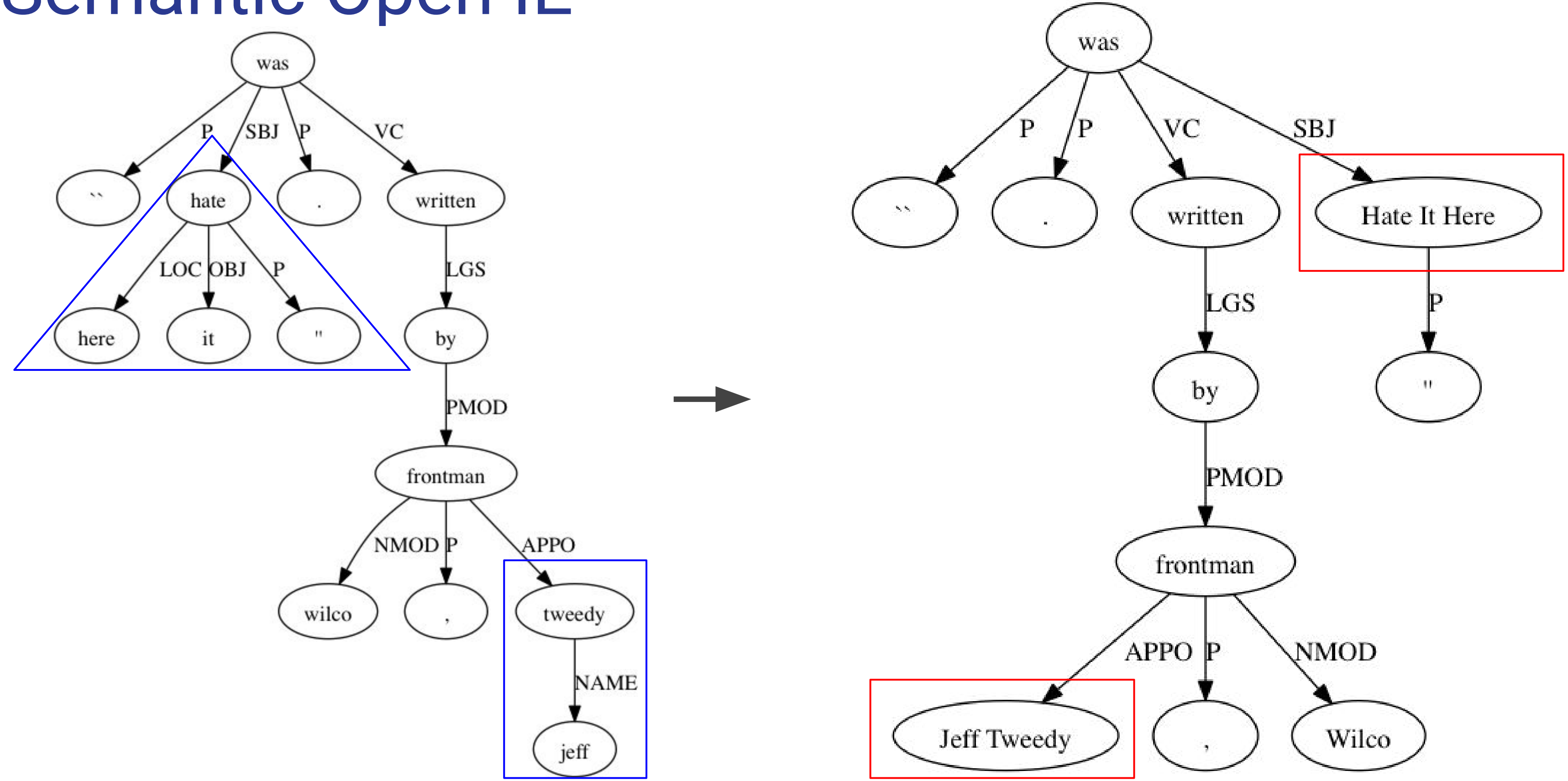
`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .

Semantic Open IE



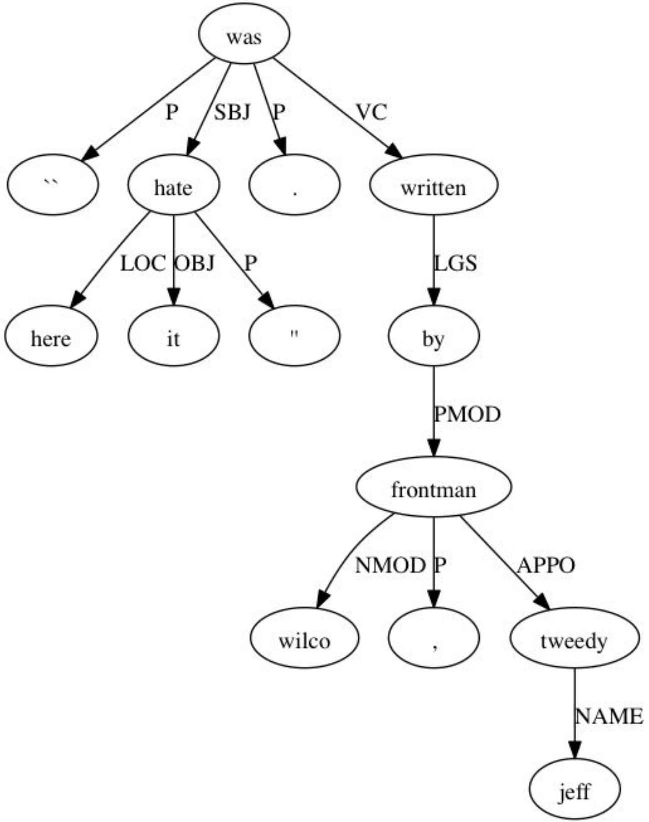
Semantic Open IE

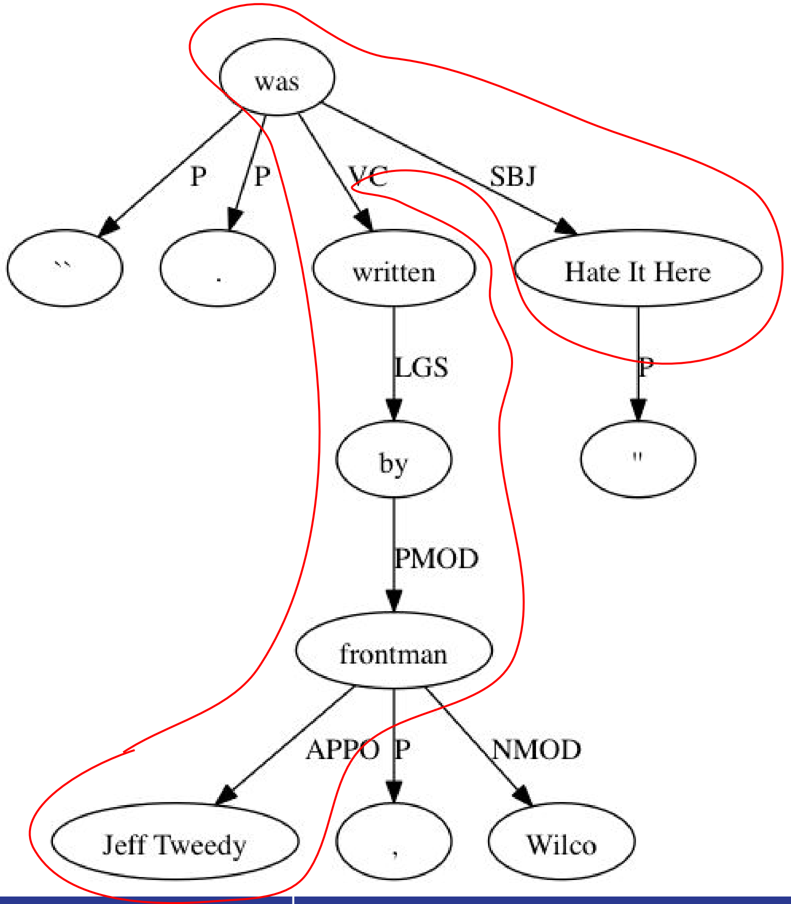
`` Hate It Here '' was written by Wilco frontman , Jeff Tweedy .



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

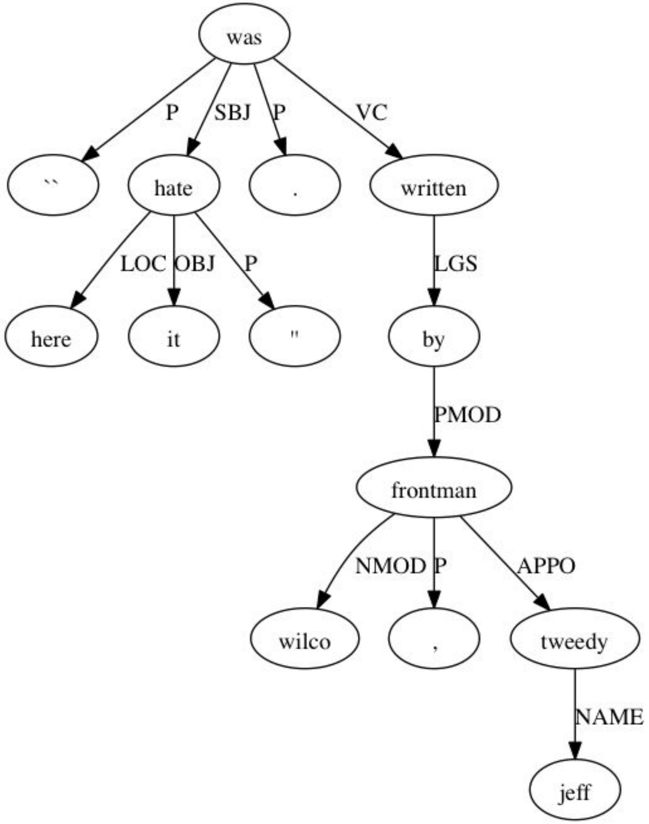
Semantic Open IE

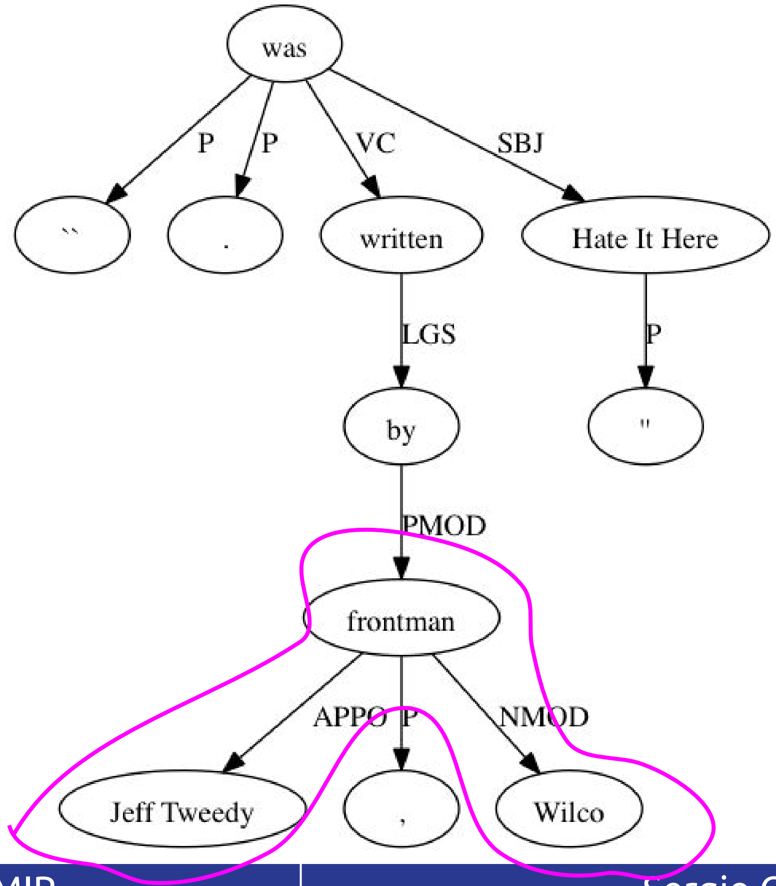




Semantic Open IE

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





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. 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**
  - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- Conclusions and Future

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

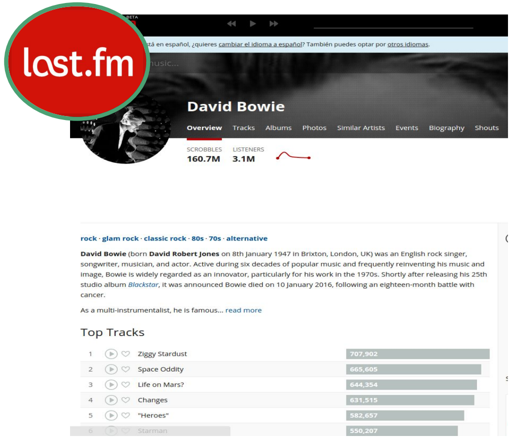
Motivation - Why you should care

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

Motivation - Why you should care

- Huge amount of music information remains implicit in unstructured texts

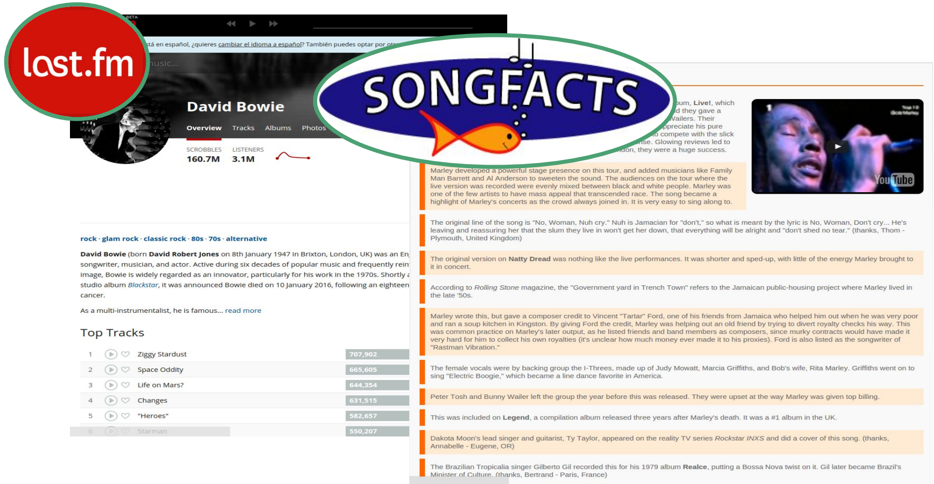
    * Artists biographies, articles, reviews, web pages, user posts.



Motivation - Why you should care

- Huge amount of music information remains implicit in unstructured texts

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



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 it there is variation?** 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!

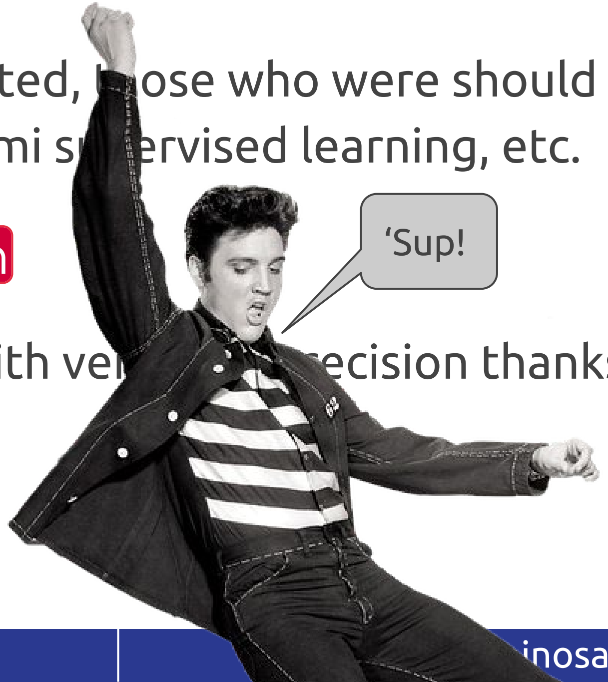
ELMD: Entity Linking in the Music Domain

- 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, [illegible]ose who were should have very high Precision. Good for propagation, semi s[illegible]ervised learning, etc.

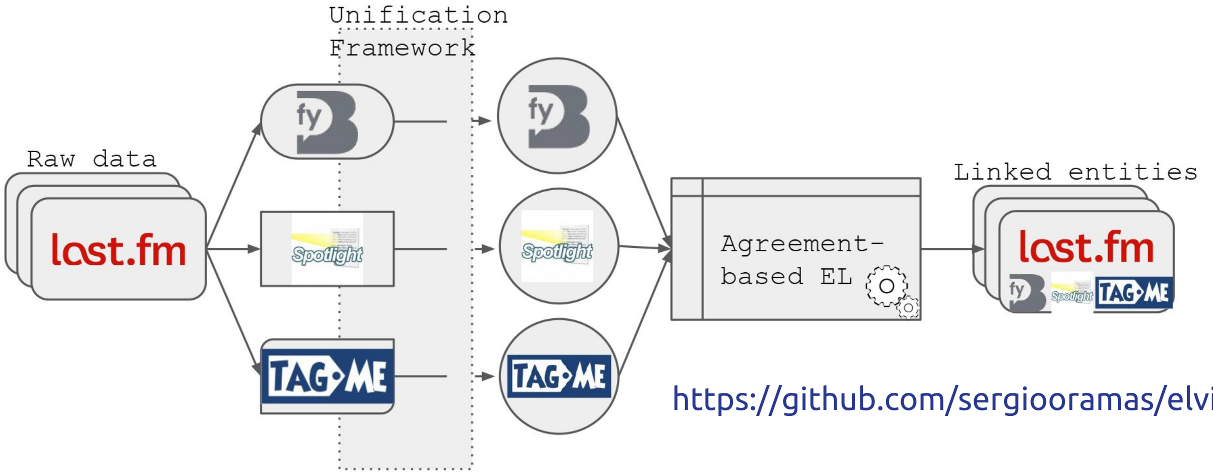
- We took advantage of artist biographies in last.fm

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



ELVIS: **E**ntity **L**inking **V**oting and **I**ntegration **S**ystem

· 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: Entity Linking in the Music Domain

	ELVIS Score	Precision	Annotations
type-equivalent	**= 3**	**0.97**	**31,180**
	>= 2	0.96	46,544
	>= 1	0.94	59,680
all	= 3	0.94	33,455
	>= 2	0.90	51,802
	>= 1	0.81	72,365

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

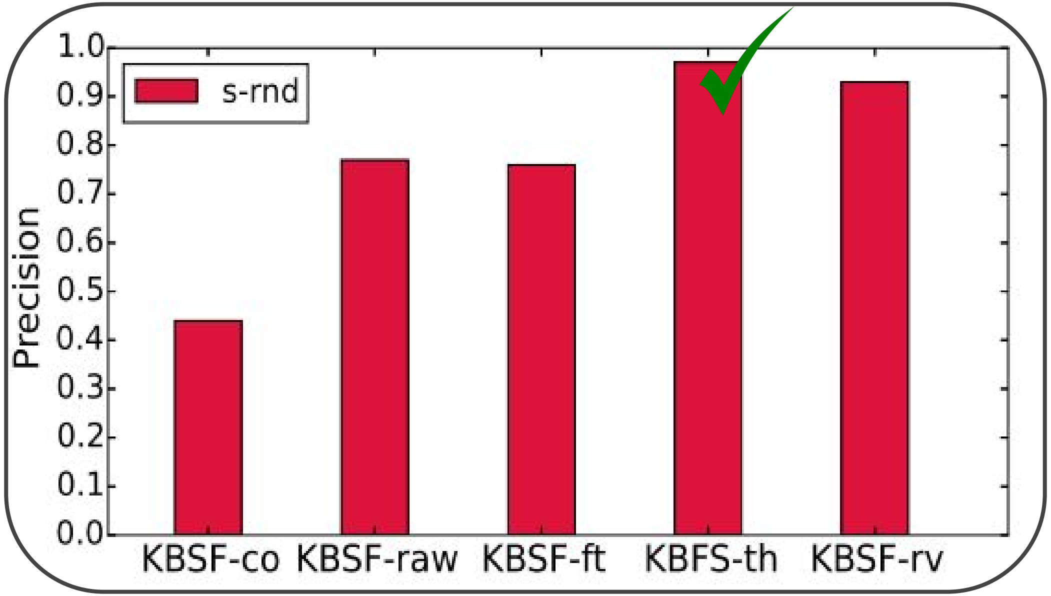
- **Relation Scoring**

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

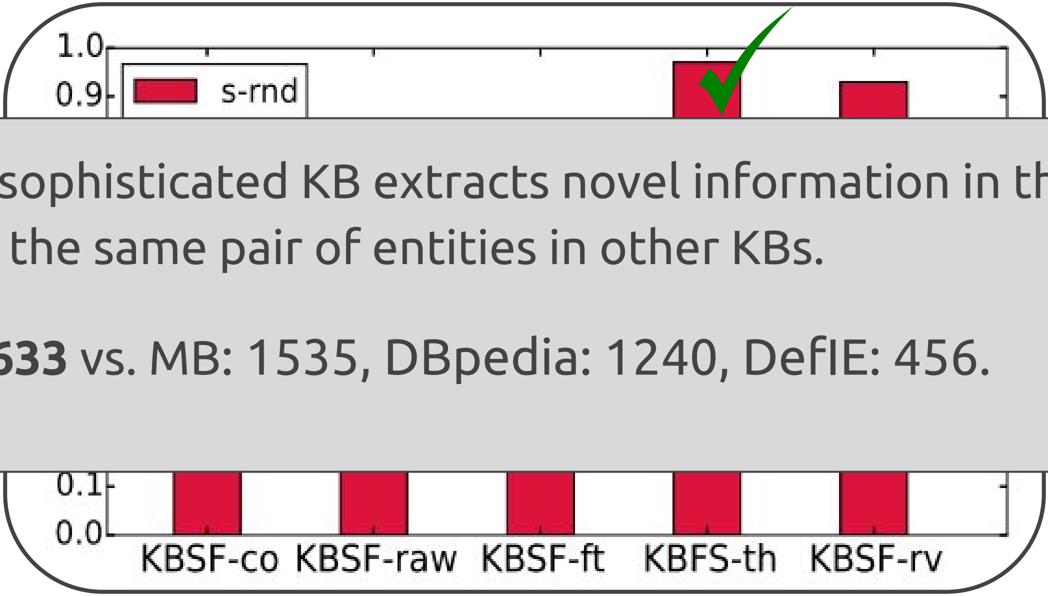
- Degree of specificity. ⇒ <*artist*$_d$, *performed_with, artist*$_r$>

- 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.

Outline

- Introduction to NLP
- Information Extraction
    - Construction of Music Knowledge Bases
    - **Semantic Enrichment of Musical Texts**
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- Conclusions and Future

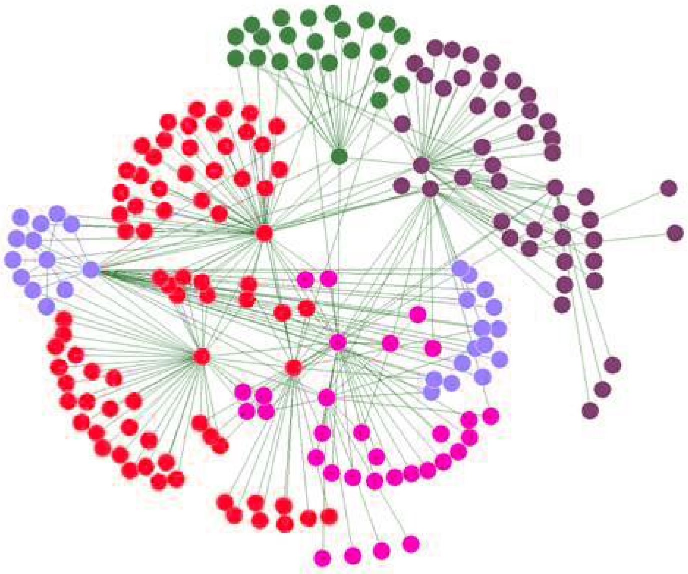
Semantic Enrichment of Musical Texts

Semantic Enrichment of Musical Texts

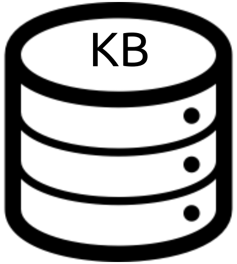
Approach: Create a Knowledge Graph and then apply graph-based methodologies or linear embeddings.

Several **types of graphs**:

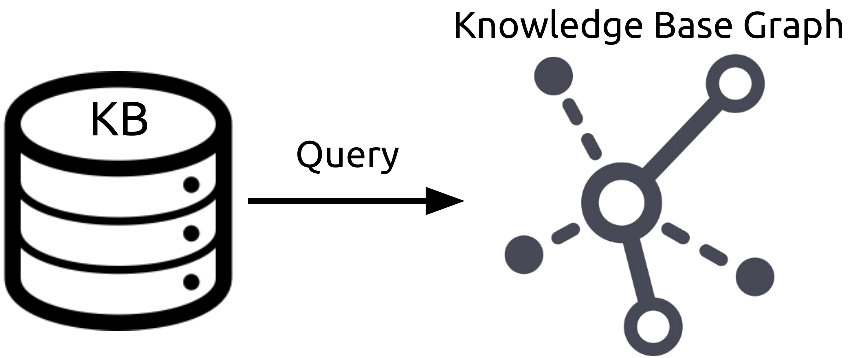
- Knowledge Base Graph
- Graph of Entities
- Semantically Enriched Graph



Knowledge Graphs



Knowledge Graphs



Knowledge Base 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 Base 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

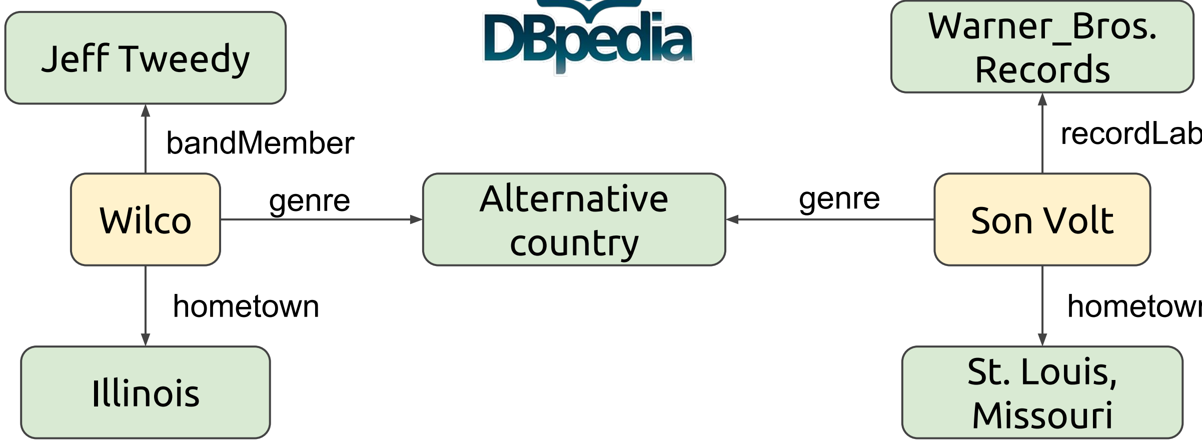
Wilco

Son Volt

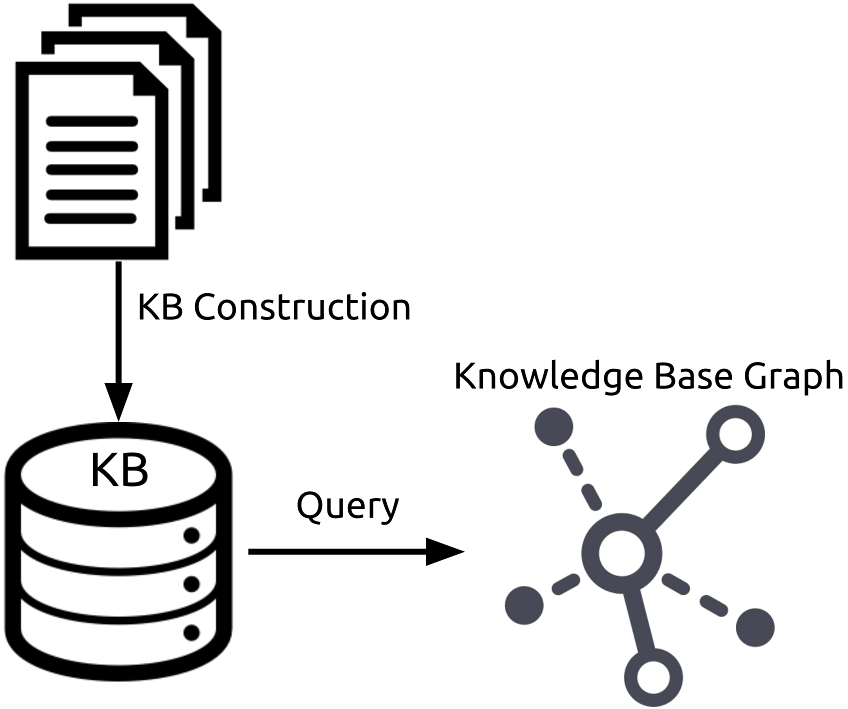
Knowledge Base Graph

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

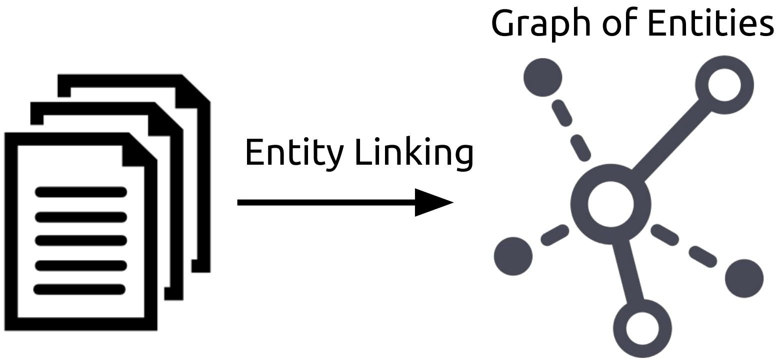
Son Volt
dbo:genre -> dbr:Alternative_country
dbo:hometown -> dbr:St._Louis,_Missouri
dbo:recordLabel -> dbr:Warner_Bros._Records



Knowledge Graphs



Knowledge 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.

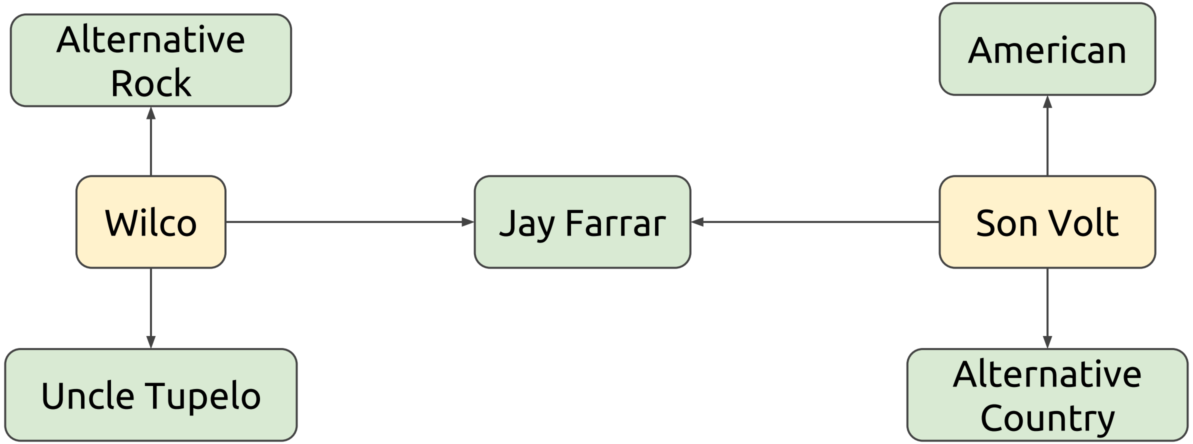
Wilco

Son Volt

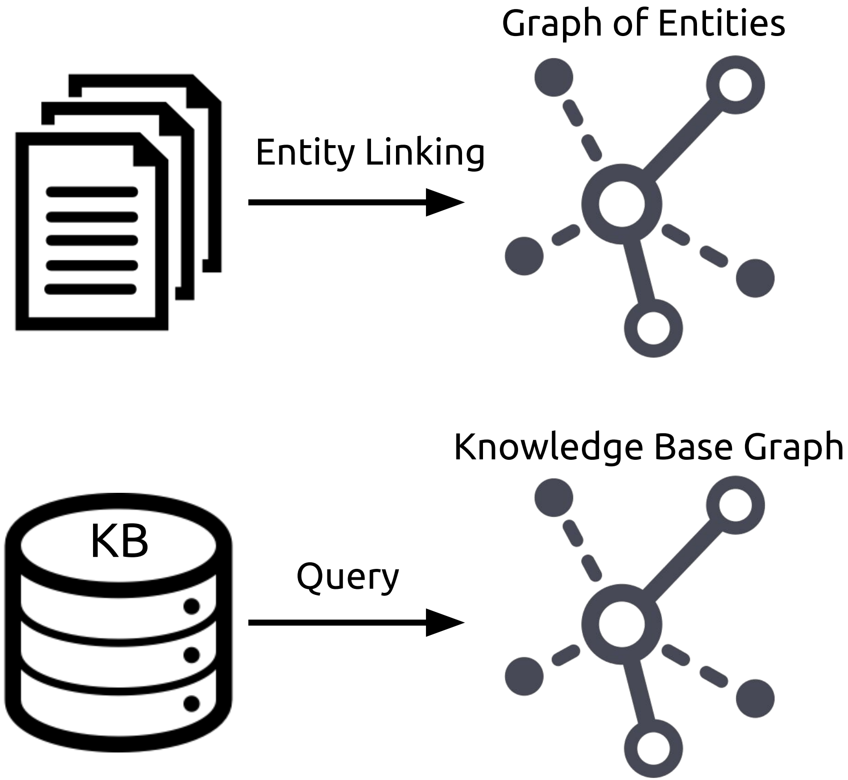
Graph of Entities

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

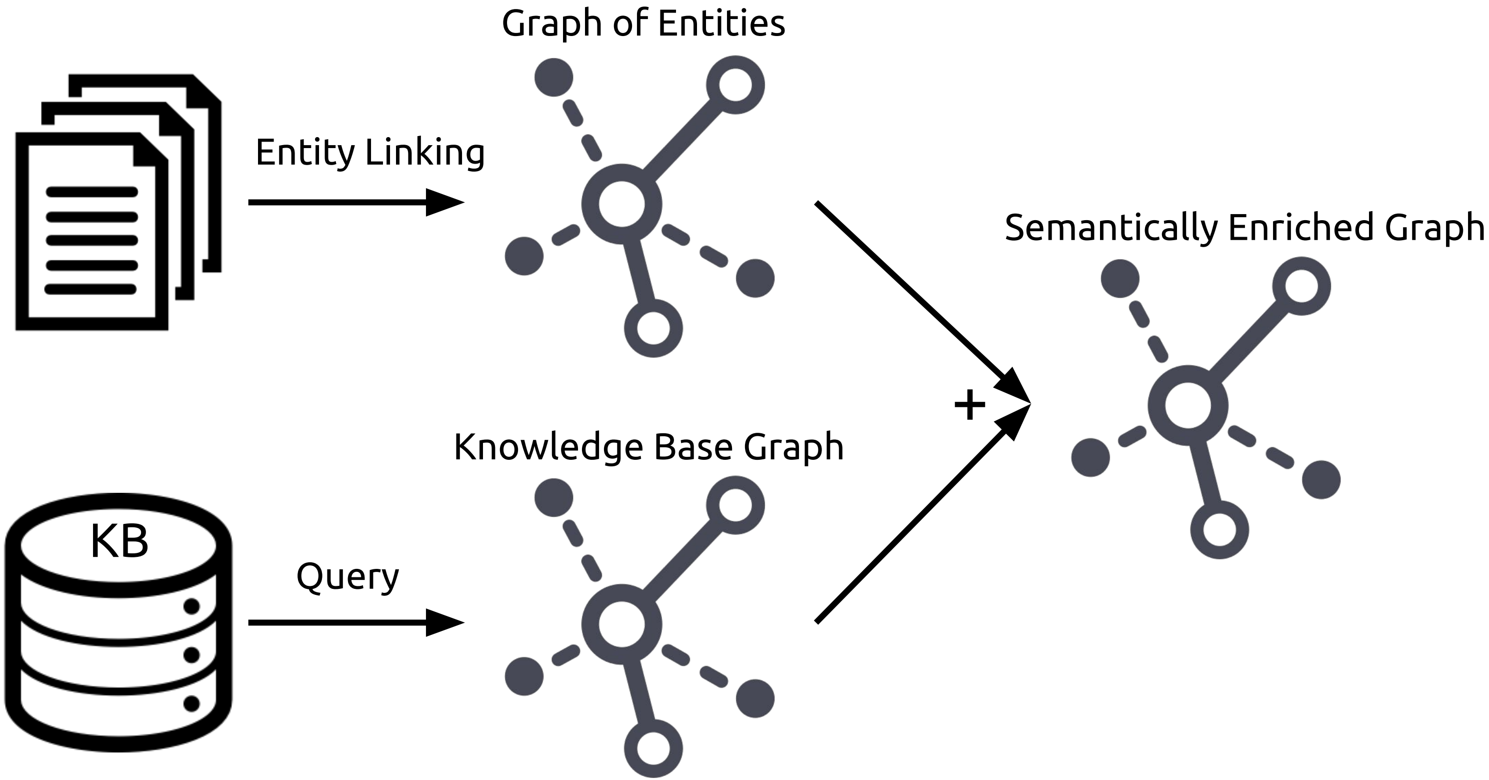
Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.



Knowledge Graphs



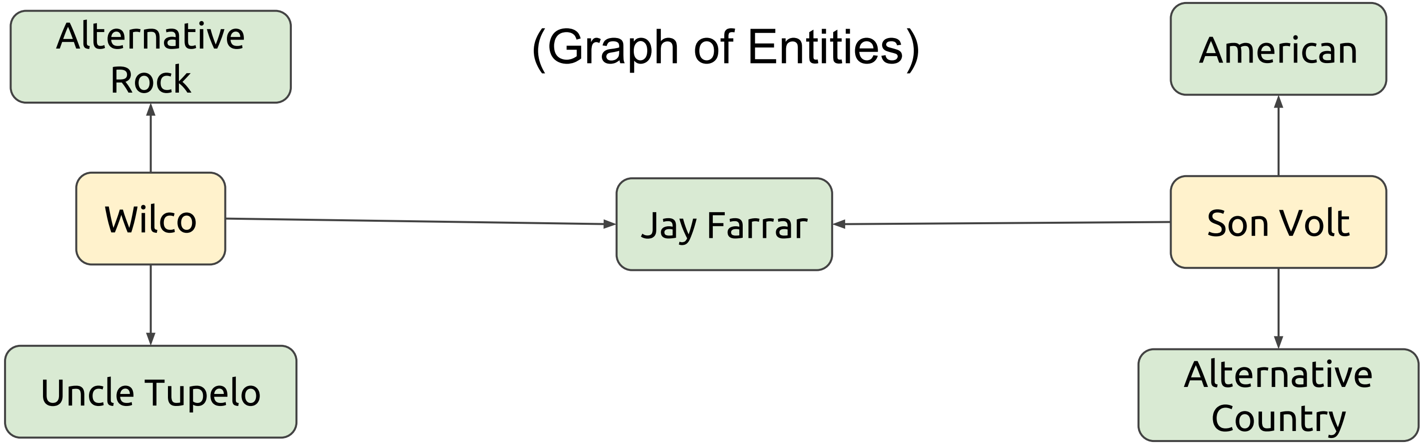
Knowledge Graphs



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.



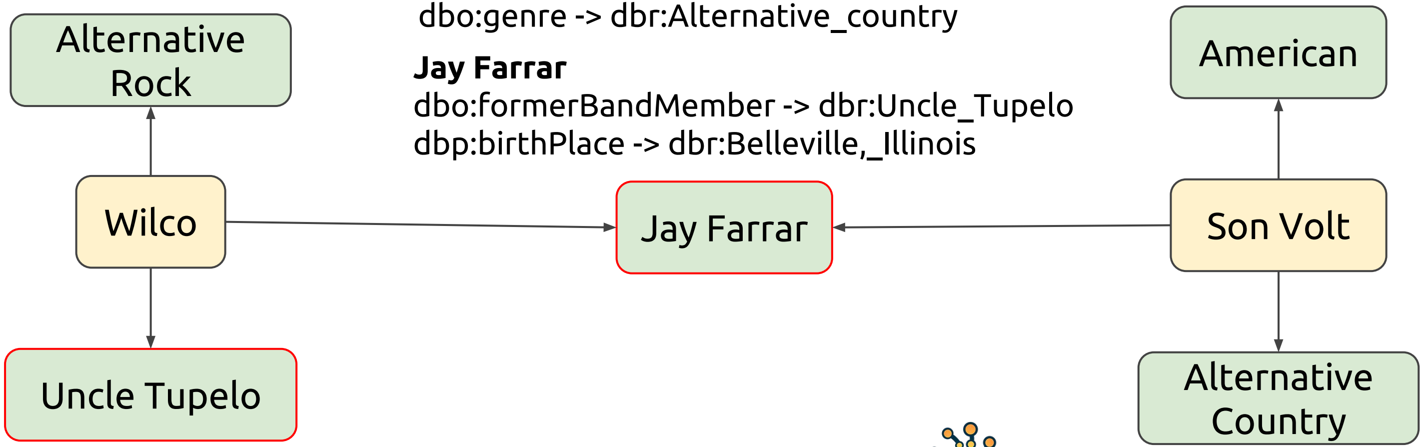
Semantically Enriched Graph

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



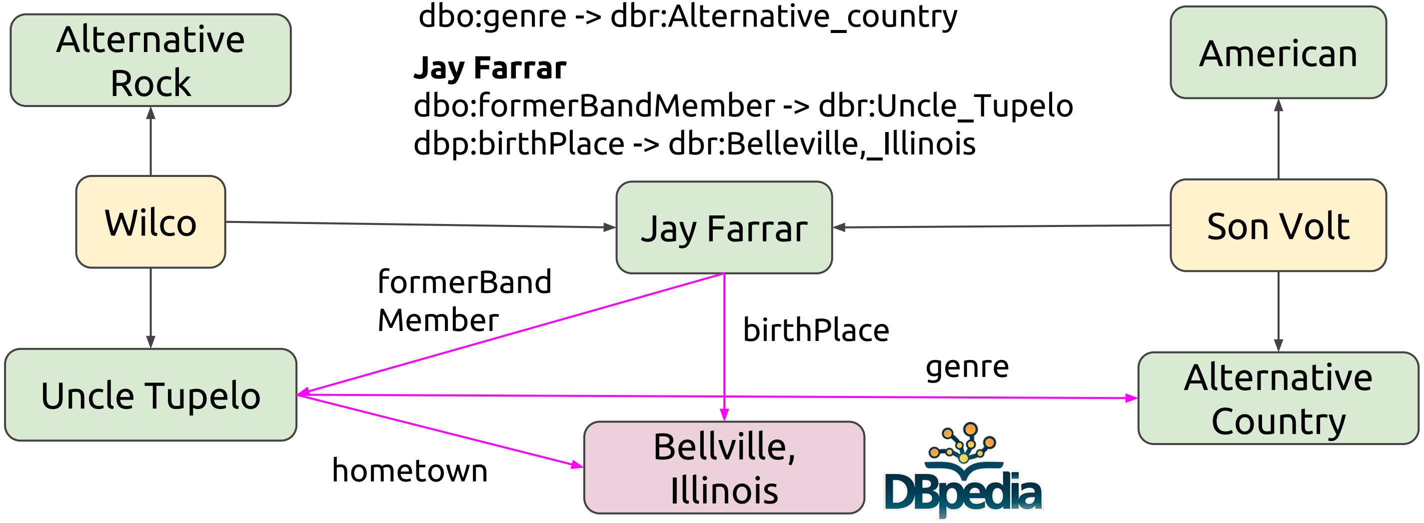
Semantically Enriched Graph

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

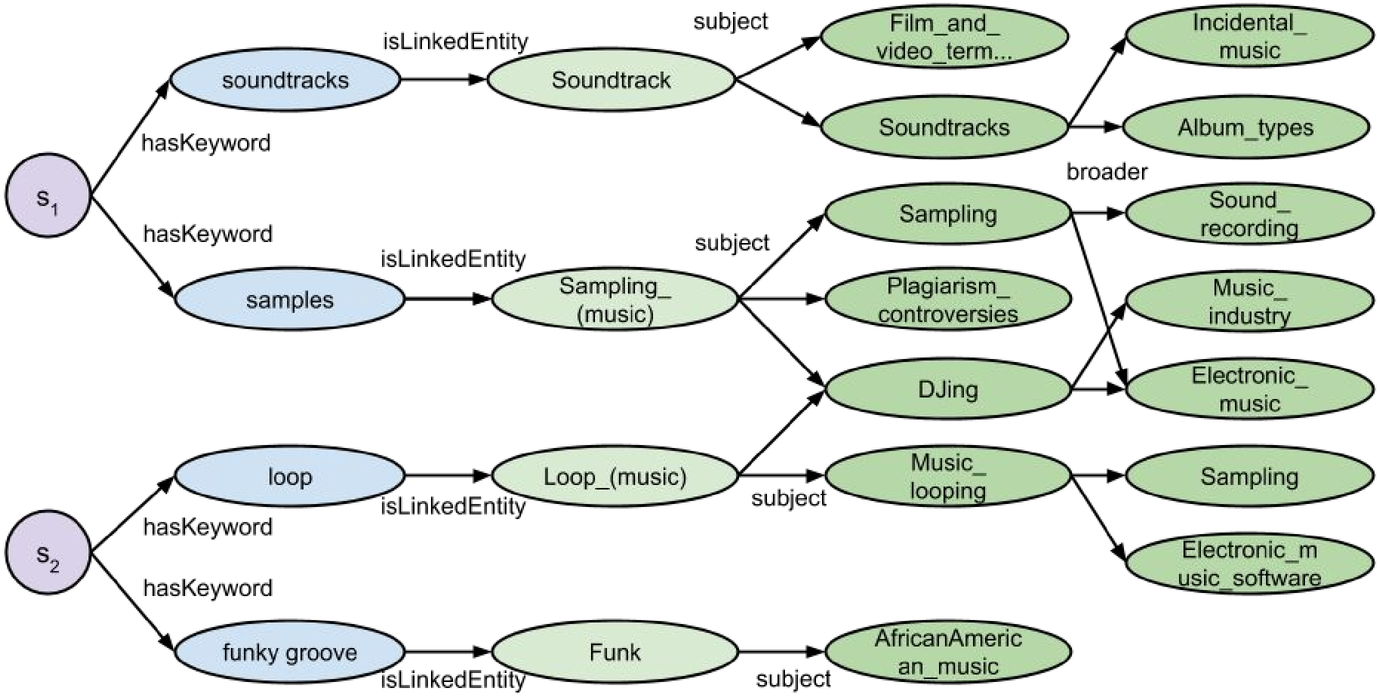
Son Volt
It is an American alternative country group, formed by Jay Farrar in 1994.

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

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



Semantically Enriched Graph



Outline

- Introduction to NLP
- Information Extraction
    - Construction of Music Knowledge Bases
    - Semantic Enrichment of Musical Texts
- **Applications in MIR**
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- Conclusions and Future

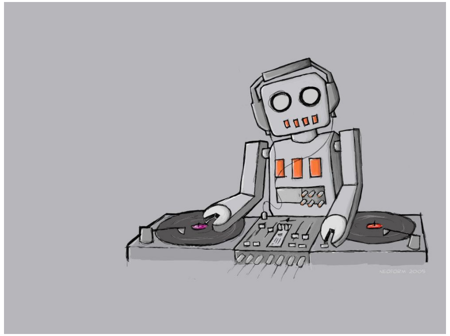
Applications in MIR

Applications

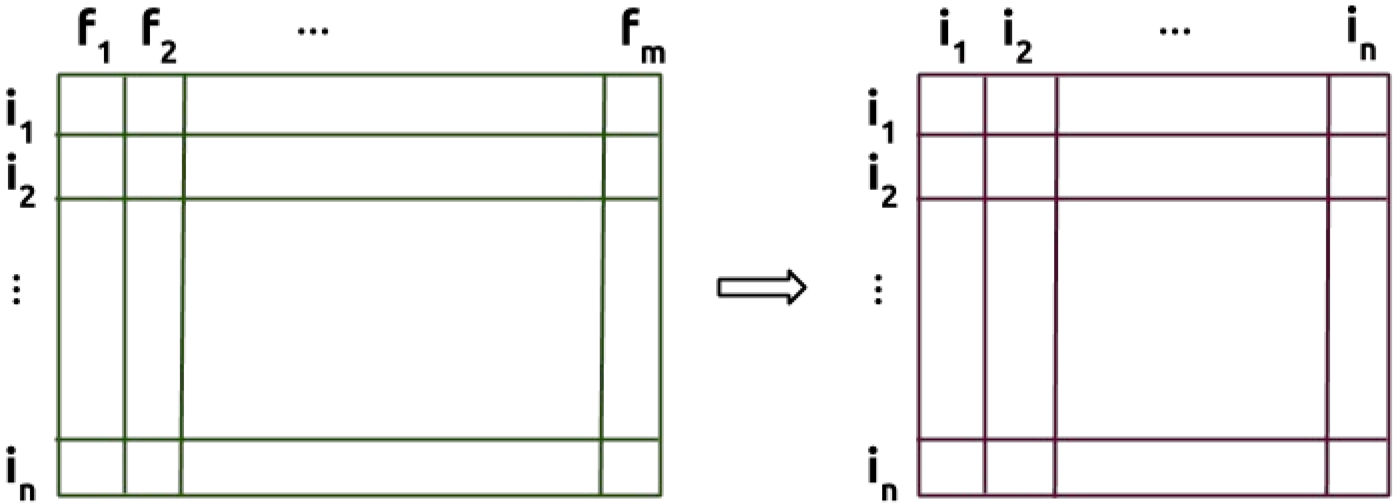
Similarity

Classification

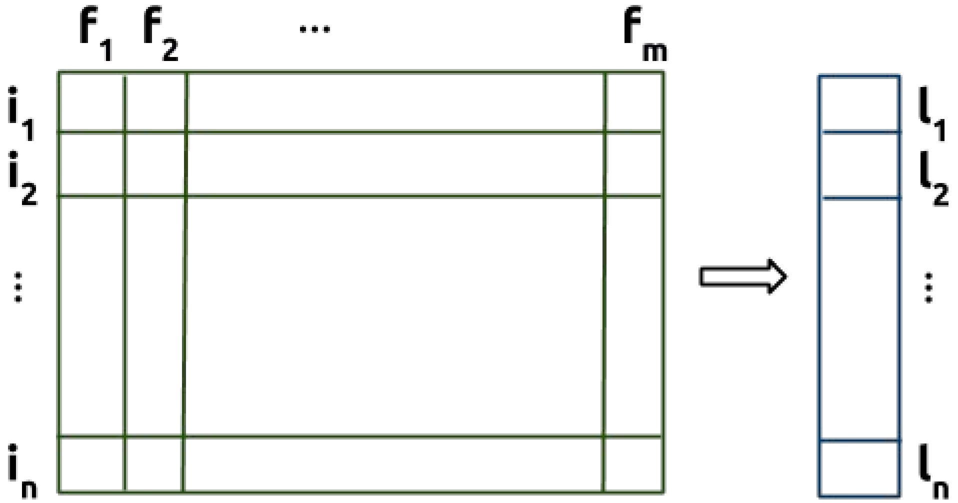
Recommendation



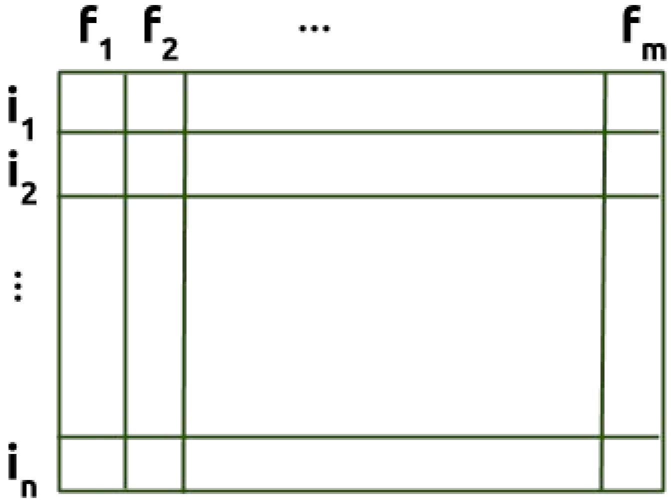
Similarity

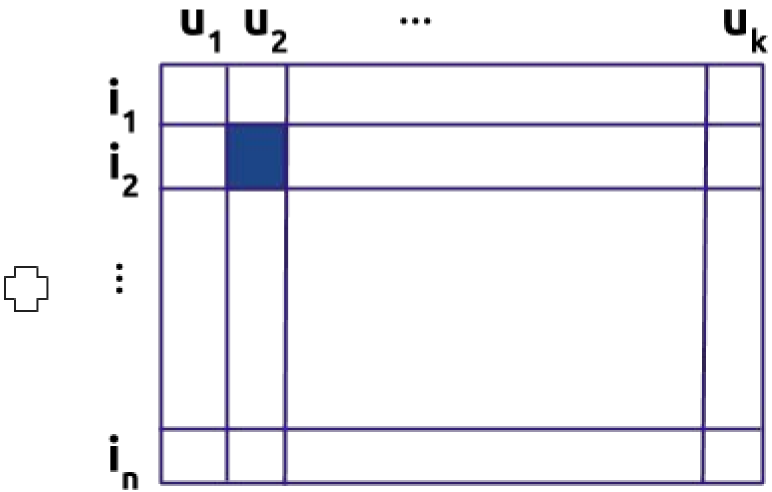


Classification



Recommendation

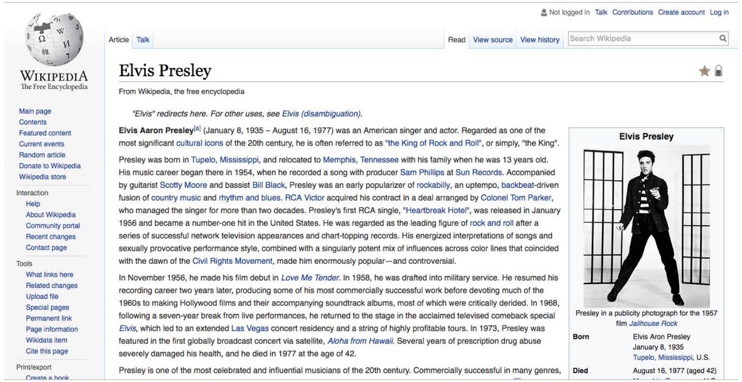


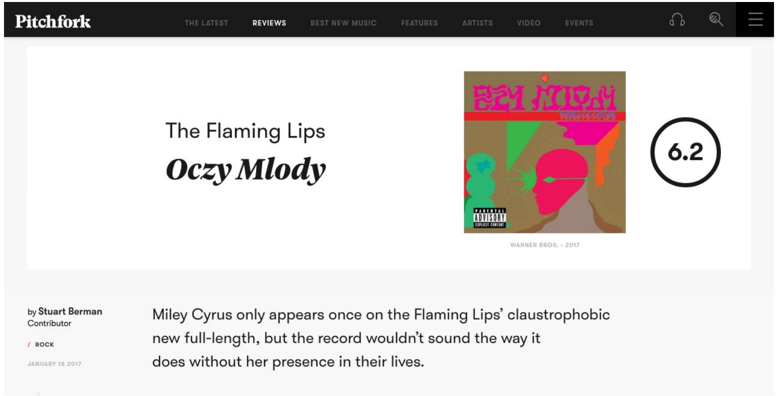


Items

Items: artist, song, sound, album

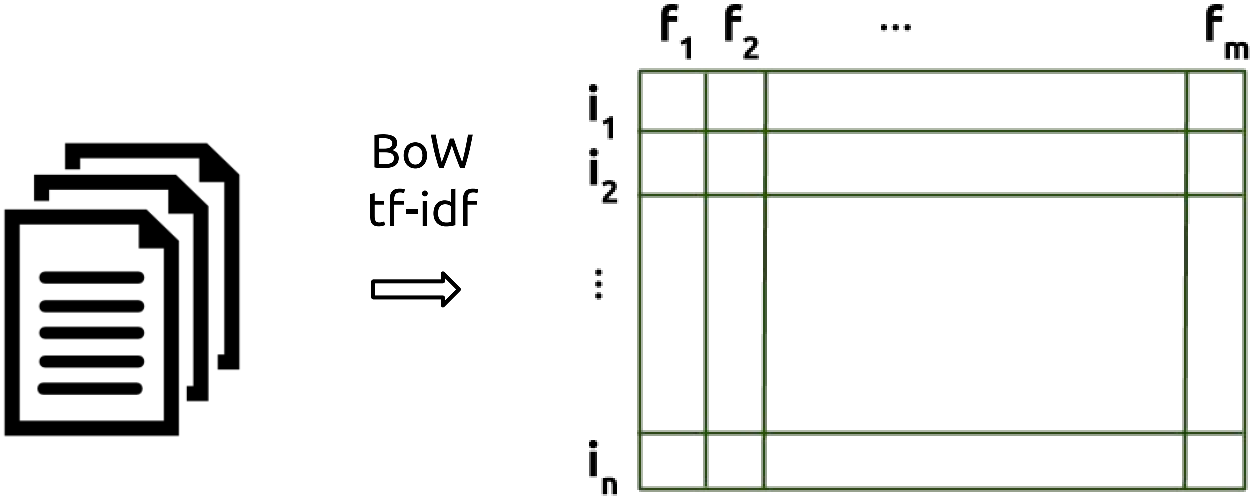
item = document



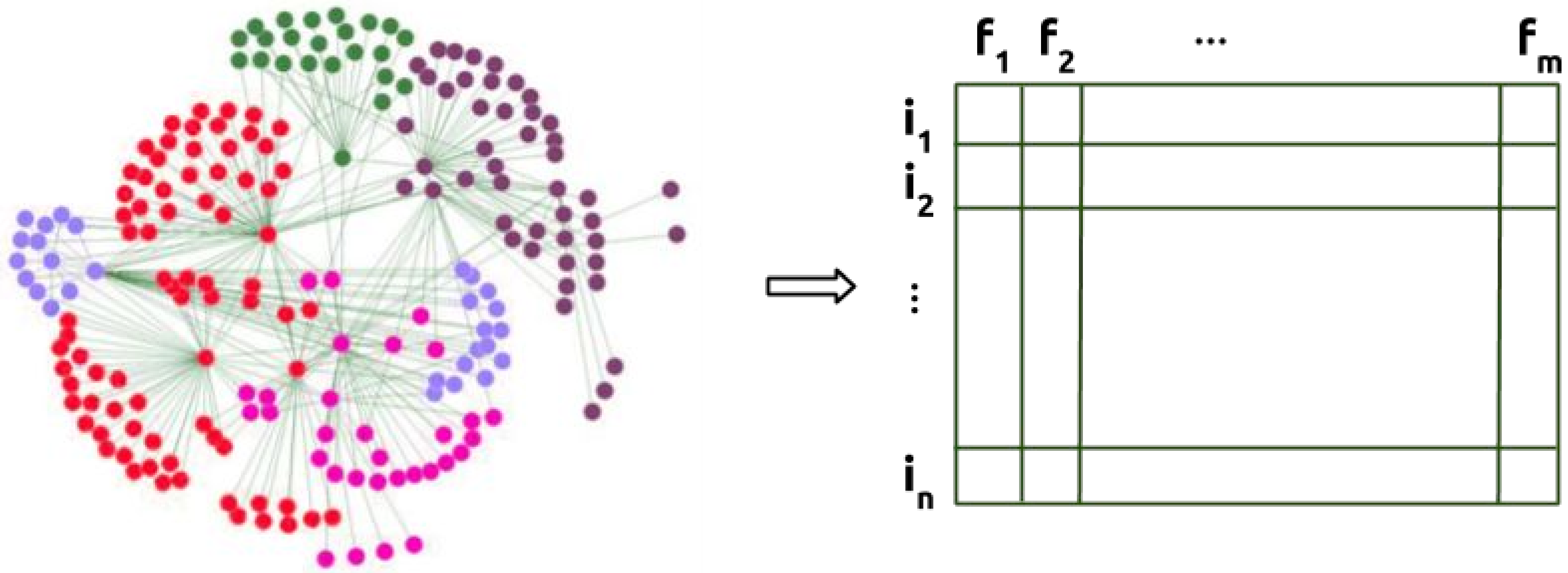


Typical Document-based approach

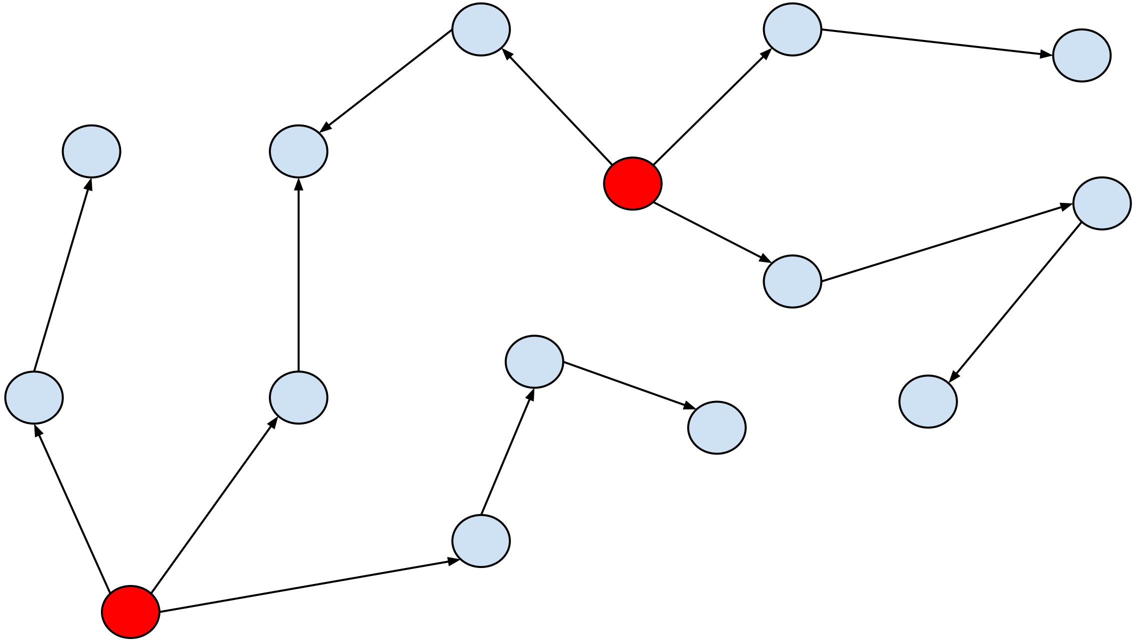
Vector Space Model



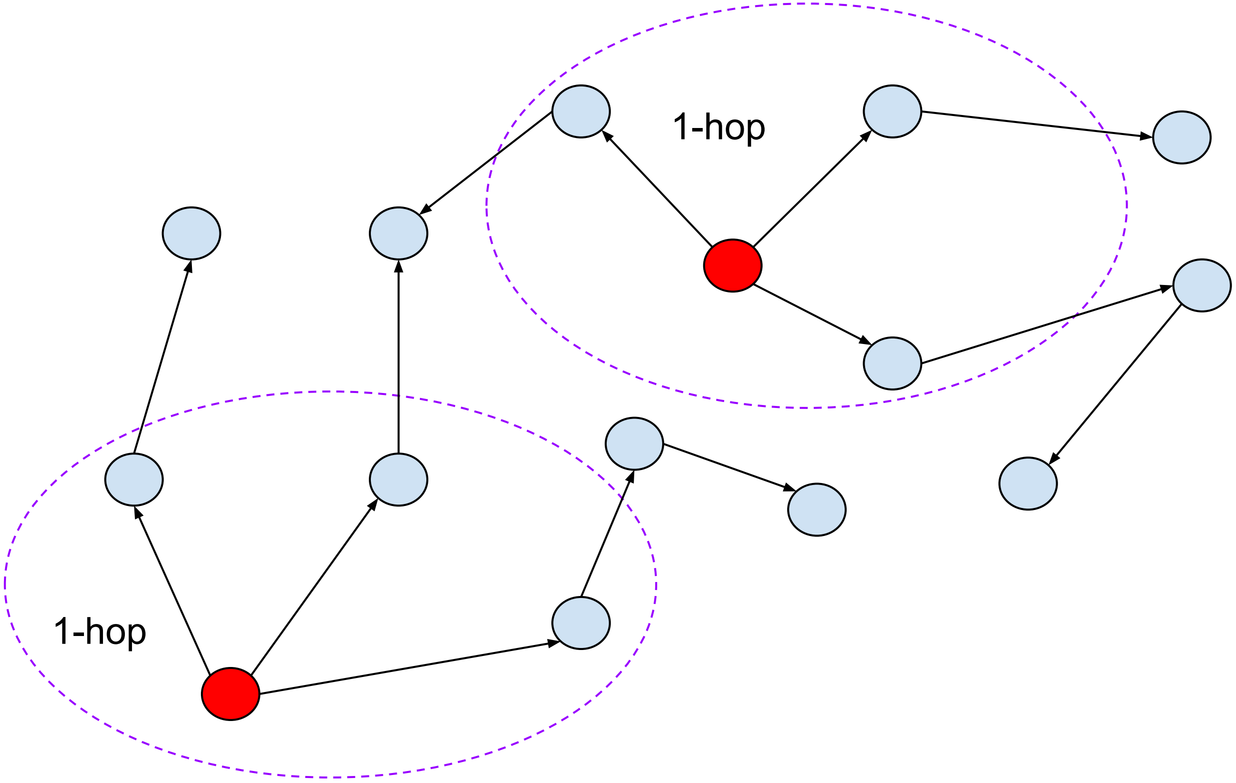
Graph Embedding



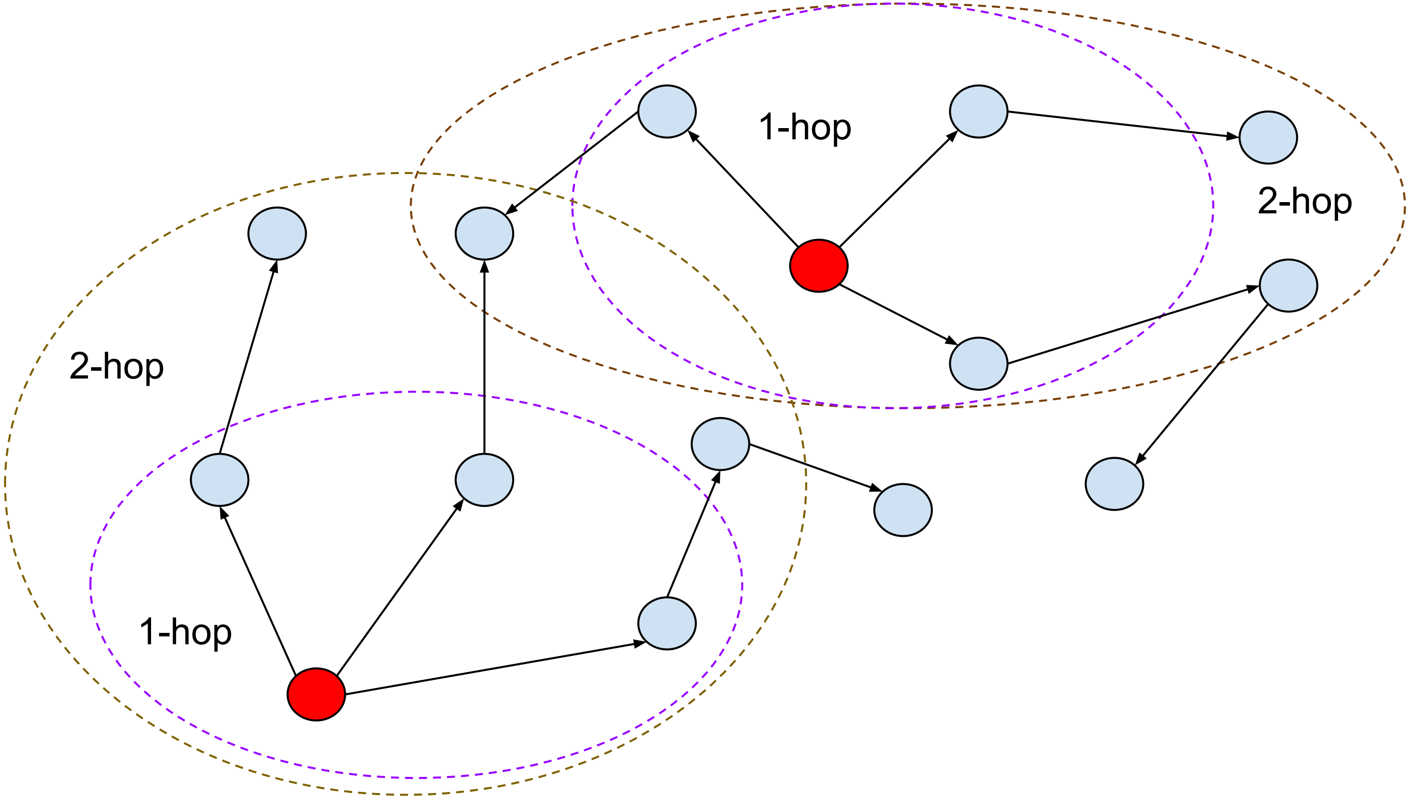
h-hop Item Neighborhood Graph



h-hop Item Neighborhood Graph



h-hop Item Neighborhood Graph



Embedding parameters

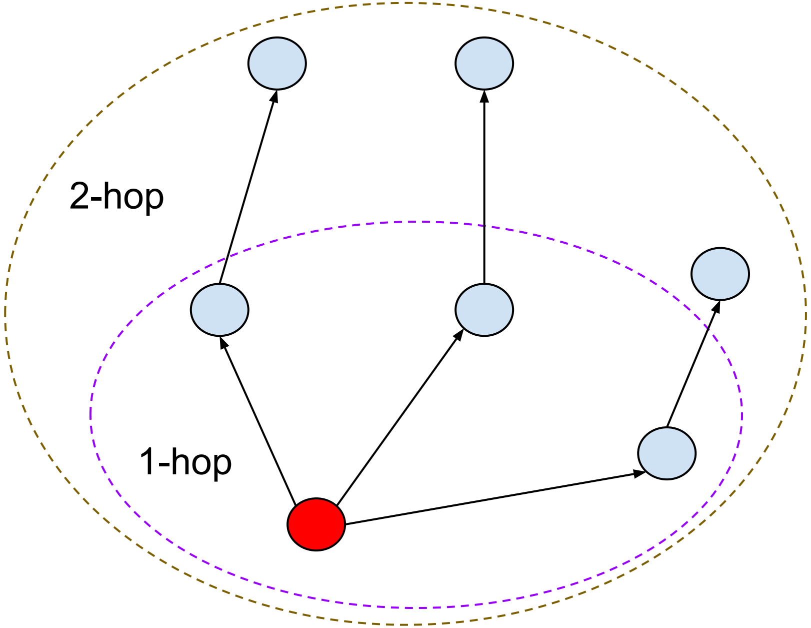
Distance to the root node

Frequency of the node inside the subgraph

Tf-idf of the node

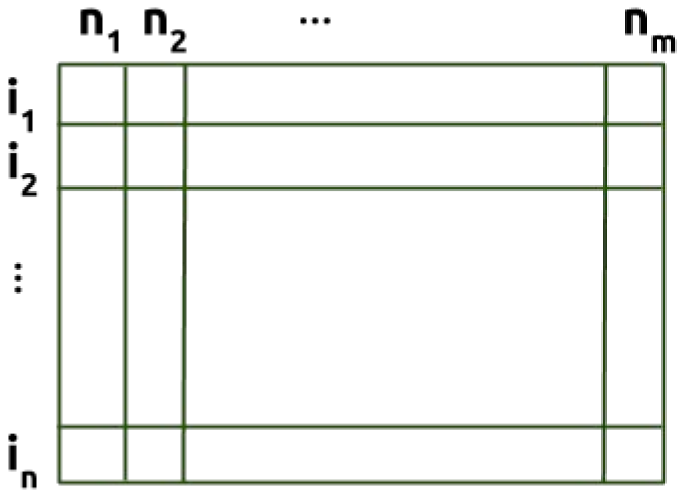
Number of in and out links

Paths: sequences of nodes from the root

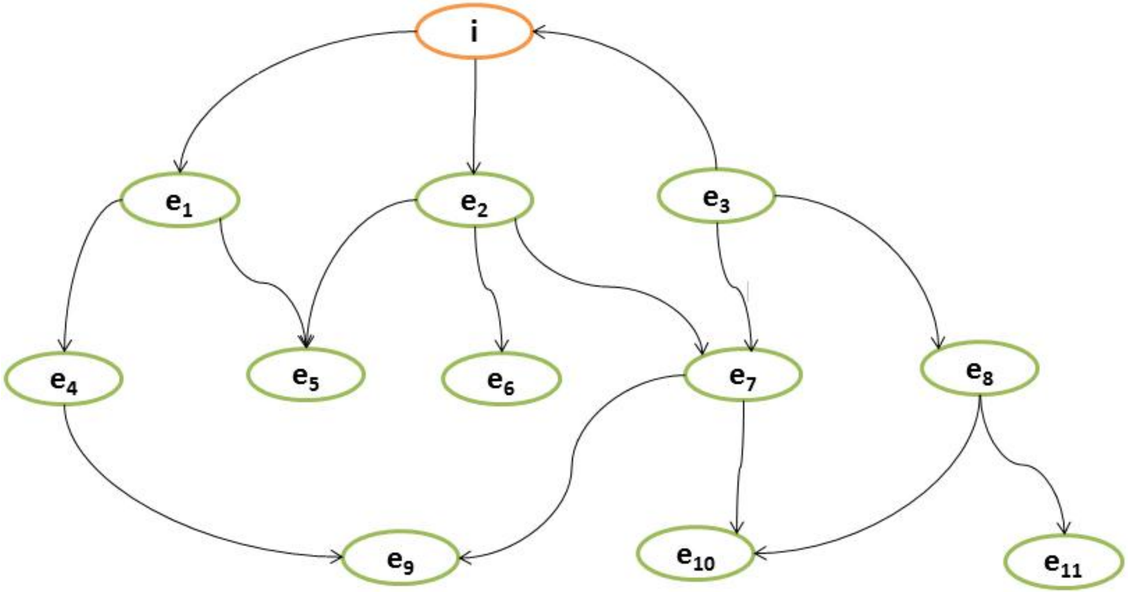


Flat Embedding

- Select h for the h-hop subgraphs
- Create a bag-of-nodes binary vector for each subgraph

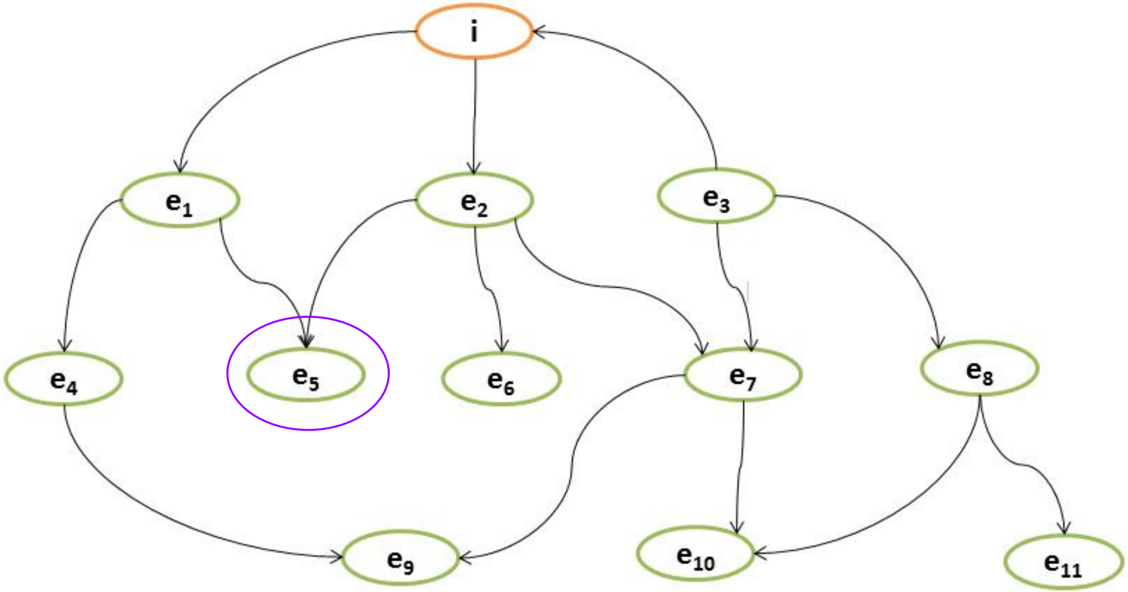


Entity-based Embedding



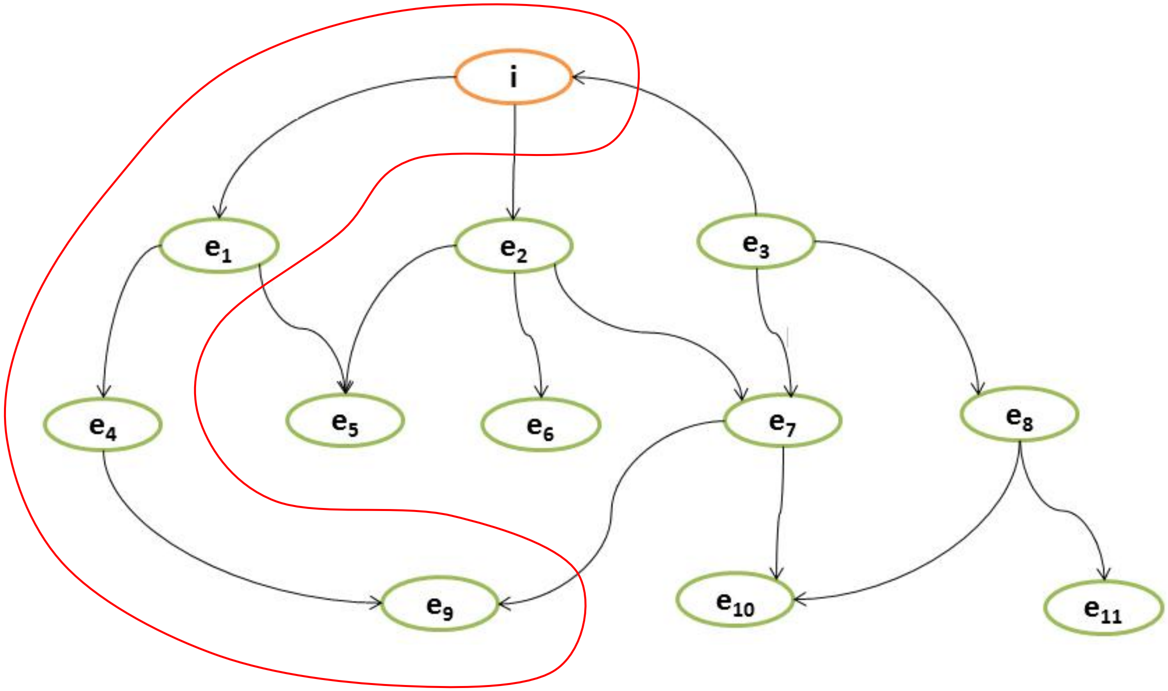
- One feature per entity
- Weight according to:
  - Distance to root
  - Number of in-links

Entity-based Embedding



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

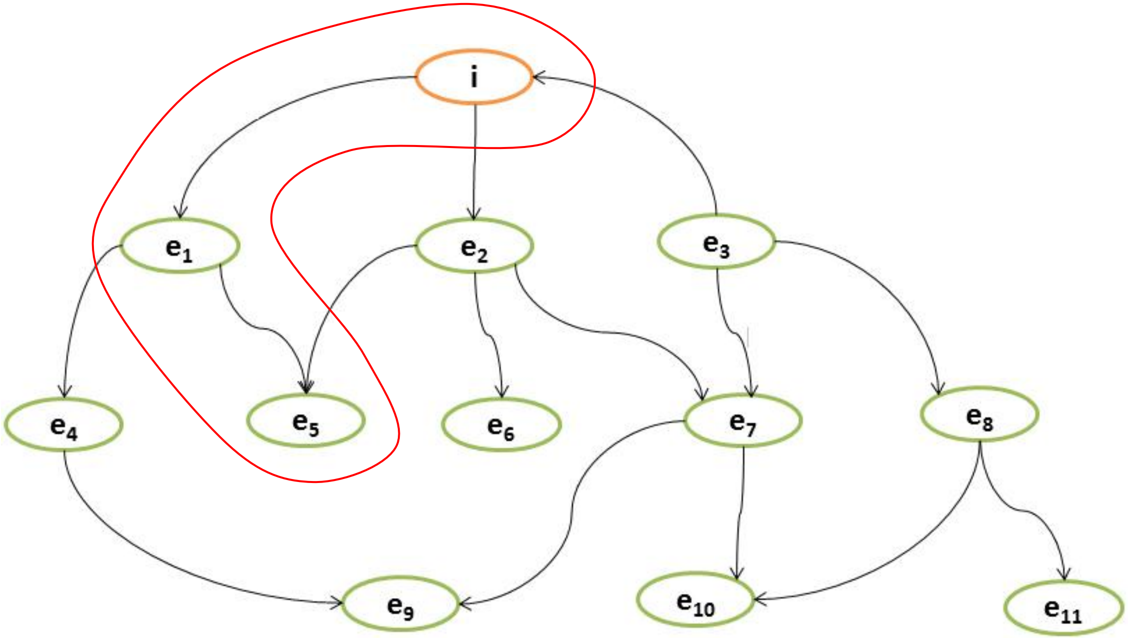
Path-based Embedding



- Path: sequence of entities

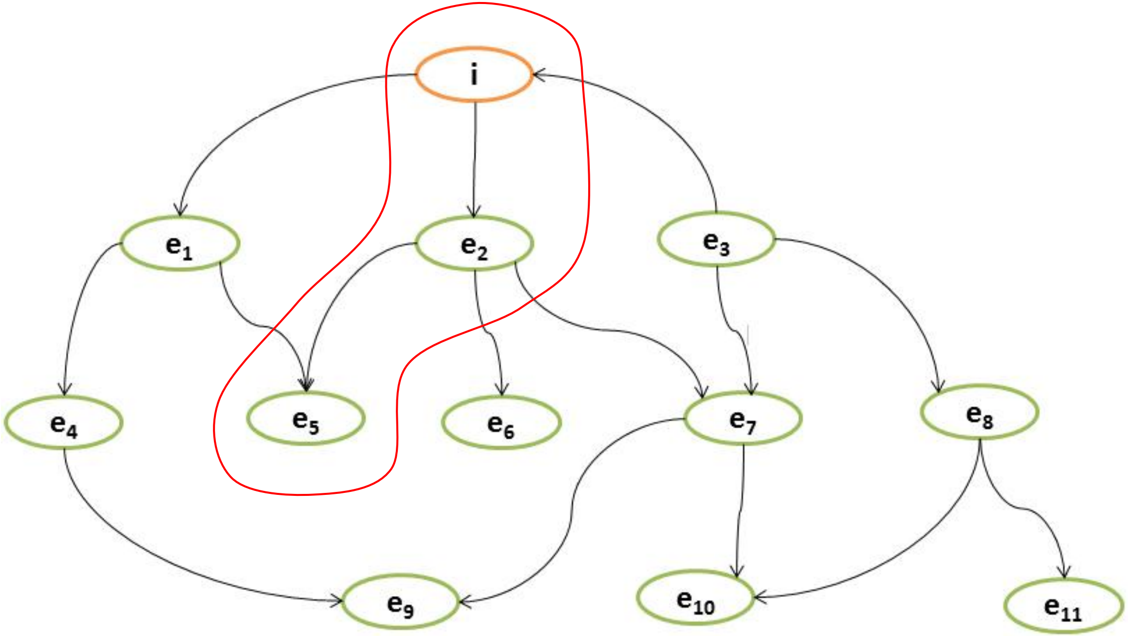
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding



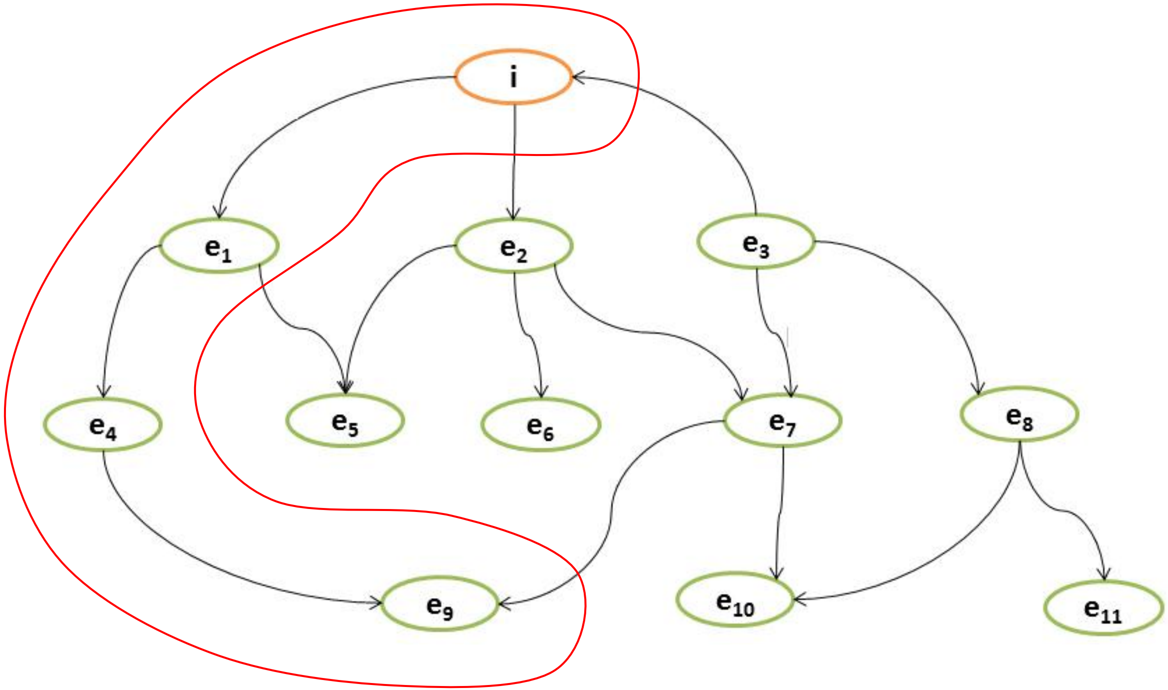
- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding



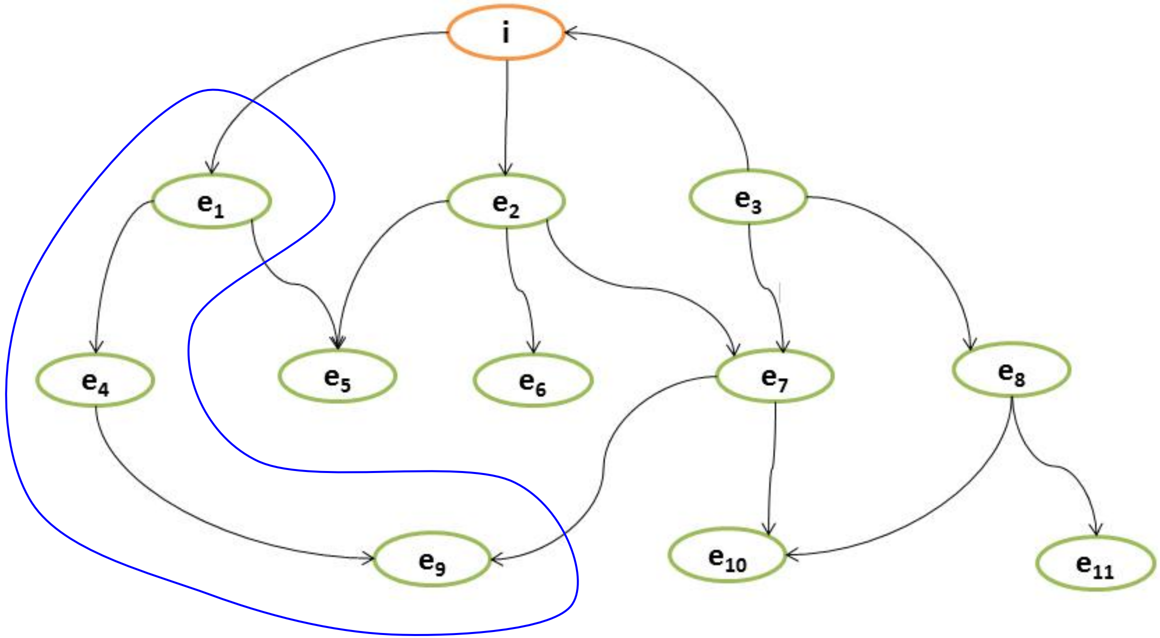
- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding



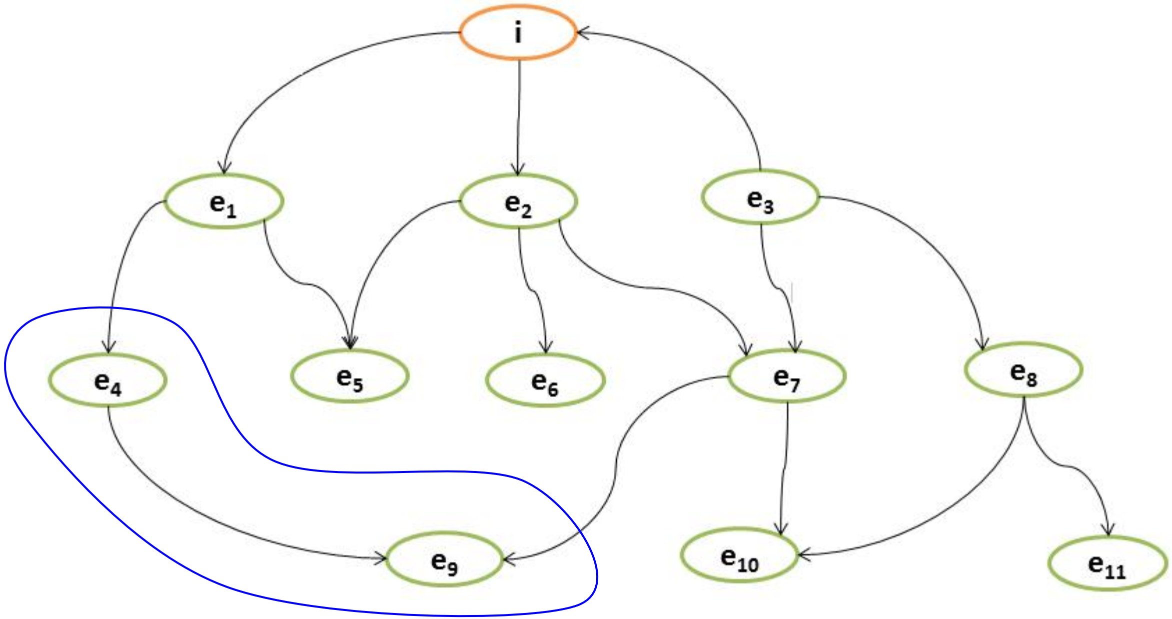
- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding



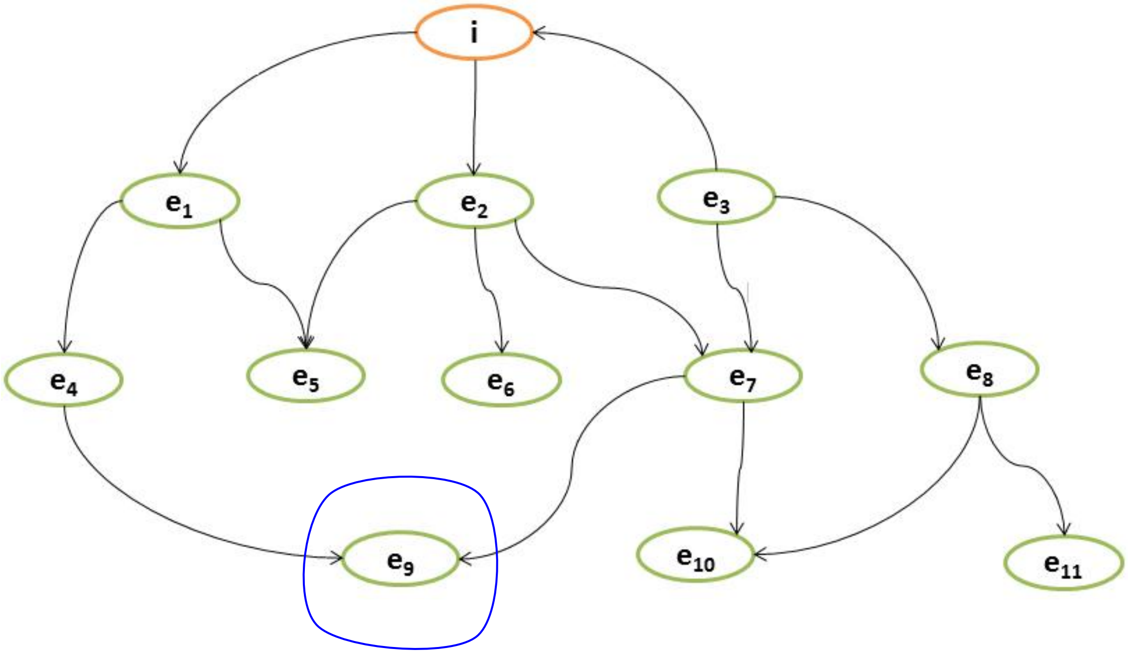
- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding



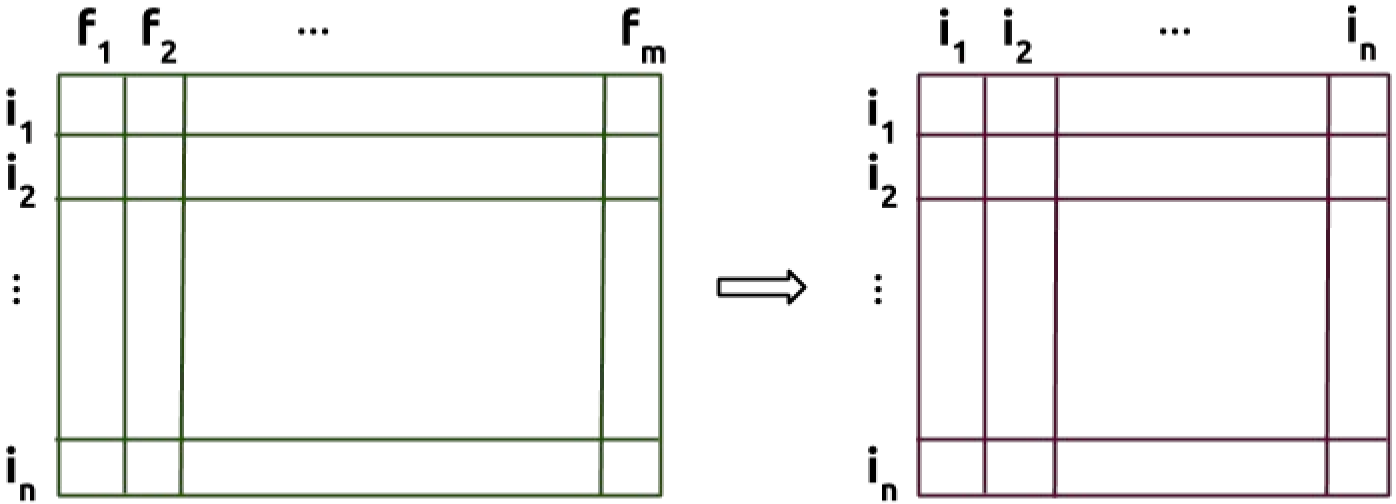
- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Path-based Embedding

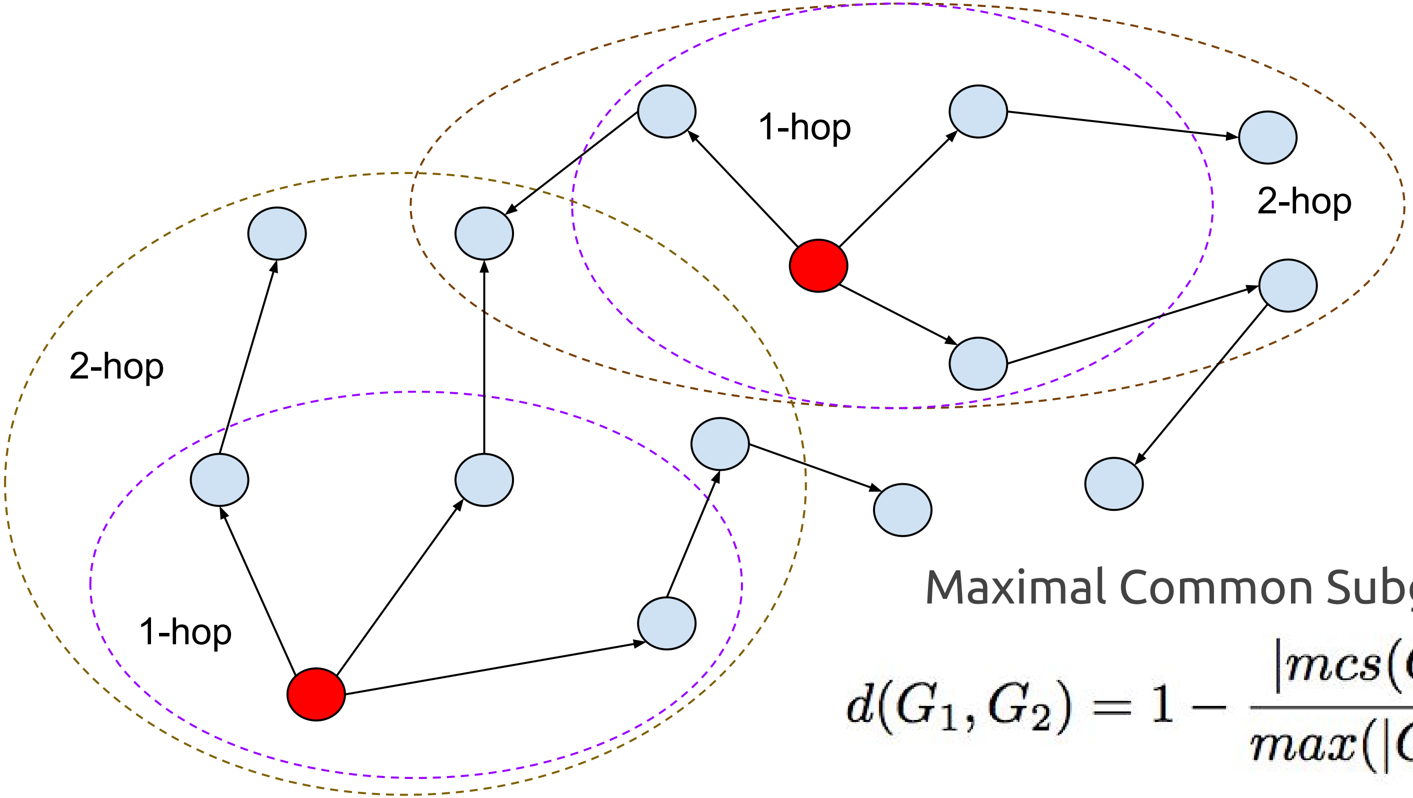


- Path: sequence of entities
- Each feature refers to several variants of paths rooted in the item node

Artist Similarity



Artist Similarity



Maximal Common Subgraph

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

Artist Similarity

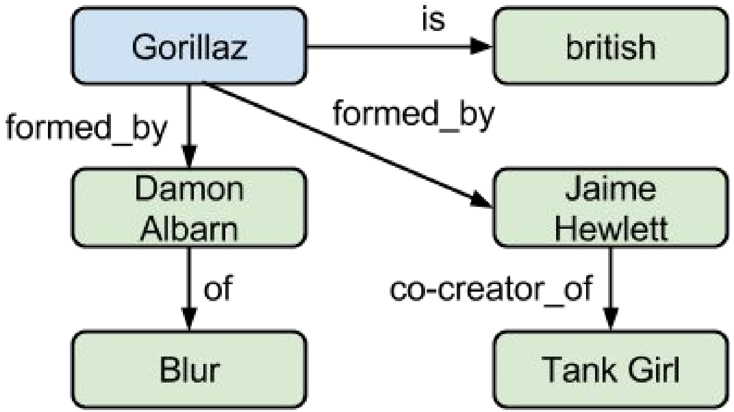
Oramas S., Sordo M., Espinosa-Anke L., & Serra X. (2015). *A Semantic-based approach for Artist Similarity.* 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

- 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

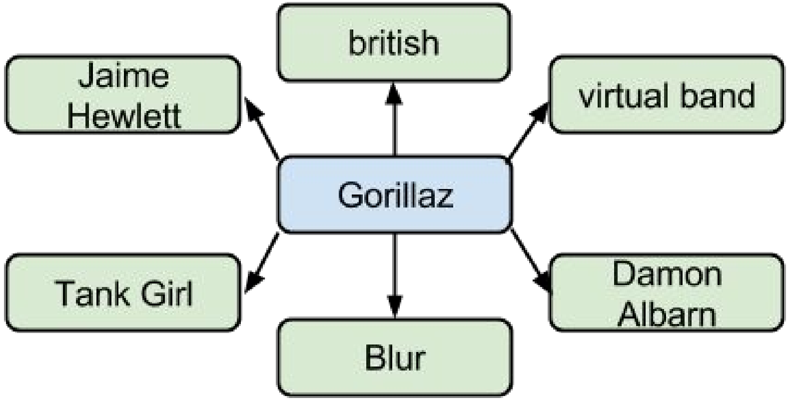
SAS dataset: http://mtg.upf.edu/download/datasets/semantic-similarity

Artist Similarity

Gorillaz are a british virtual band formed in 1998 by Damon Albarn of Blur, and Jamie Hewlett, co-creator of the comic book Tank Girl.



Extracted Knowledge Base Graph

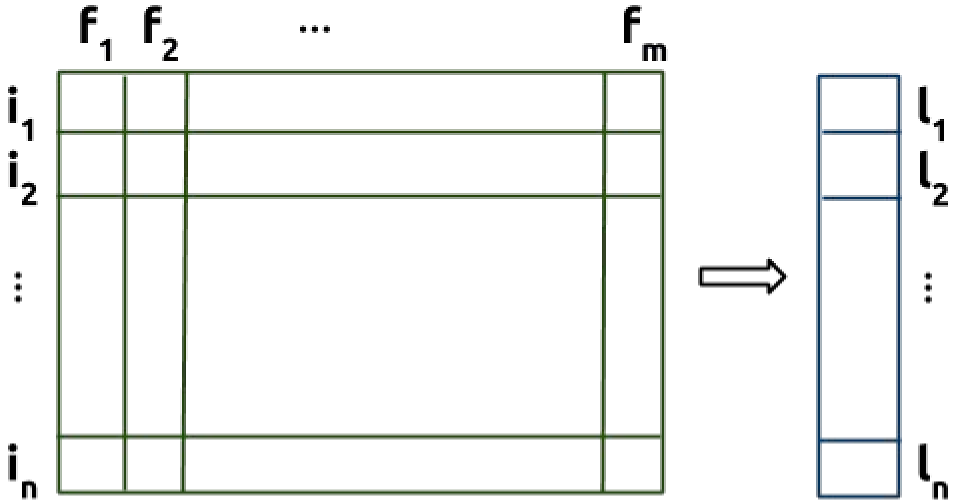


Graph of Entities

Artist Similarity

Approach	P@5
Text-based approach (BoW)	0.090
Extracted KB Graph	0.055
Graph of Entities	0.136
Semantically Enriched Graph	0.160

Genre Classification



Genre Classification

MARD (Multimodal Album Reviews Dataset):

New dataset of album customer reviews from:

Amazon + MusicBrainz + AcousticBrainz

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.

Genre Classification

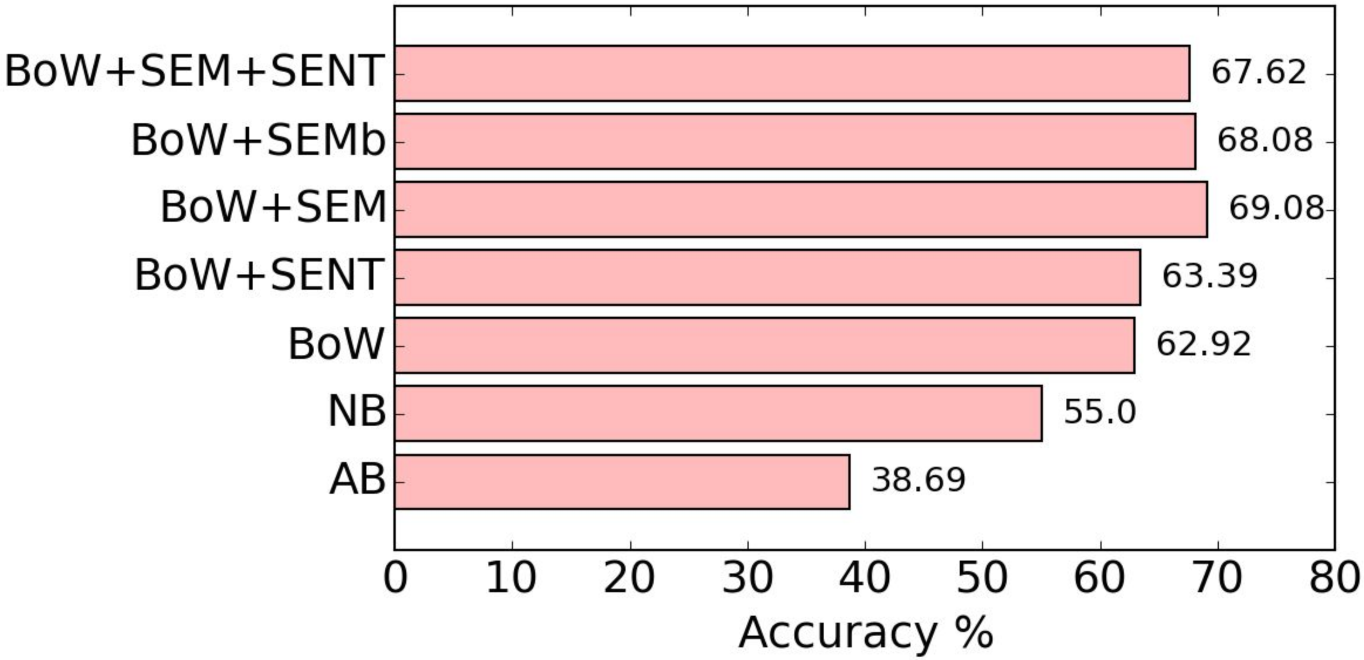
Features:

- **Textual**: BoW uni-grams and bi-grams
- **Semantic**: Entities and Wikipedia categories (Entity Linking), flat embedding
- **Sentiment**: positiveness ratio, emotion ratio, average emotion strength
- **Acoustic**: low-level descriptors (loudness, dynamics, spectral shape, etc.)

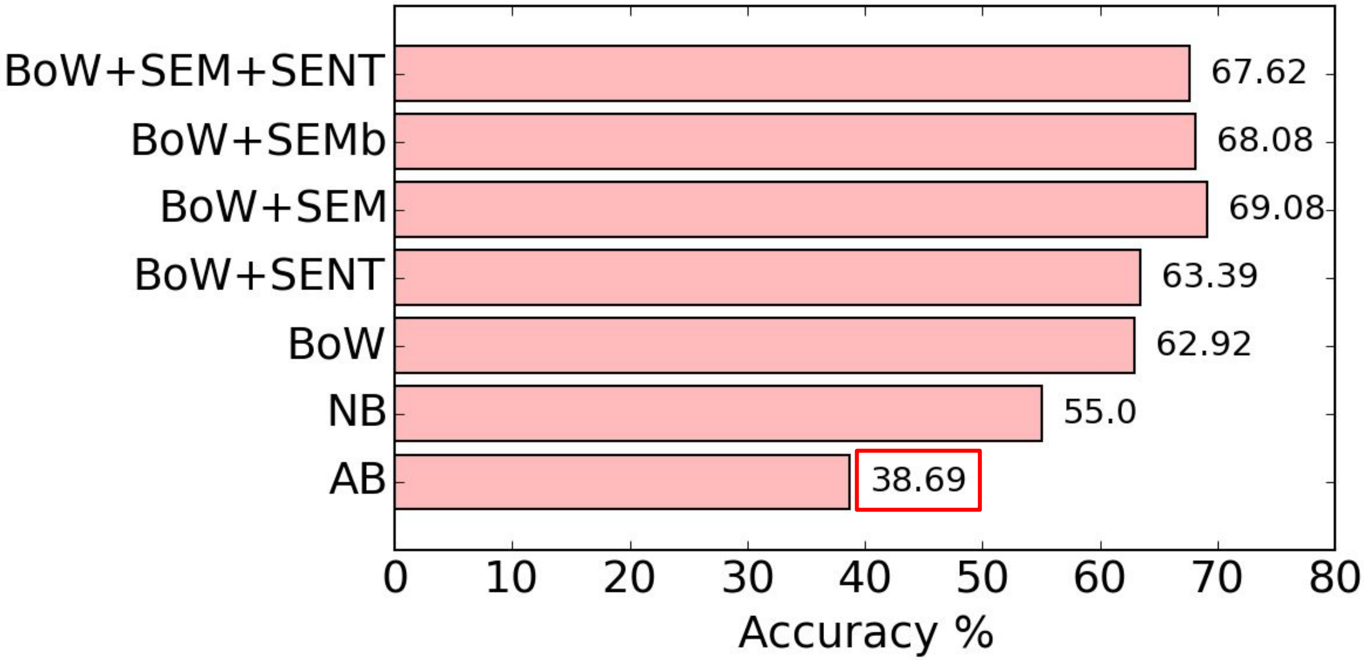
SVM classifier 5-fold cross validation, 1300 albums, 13 genres

https://github.com/sergiooramas/music-genre-classification

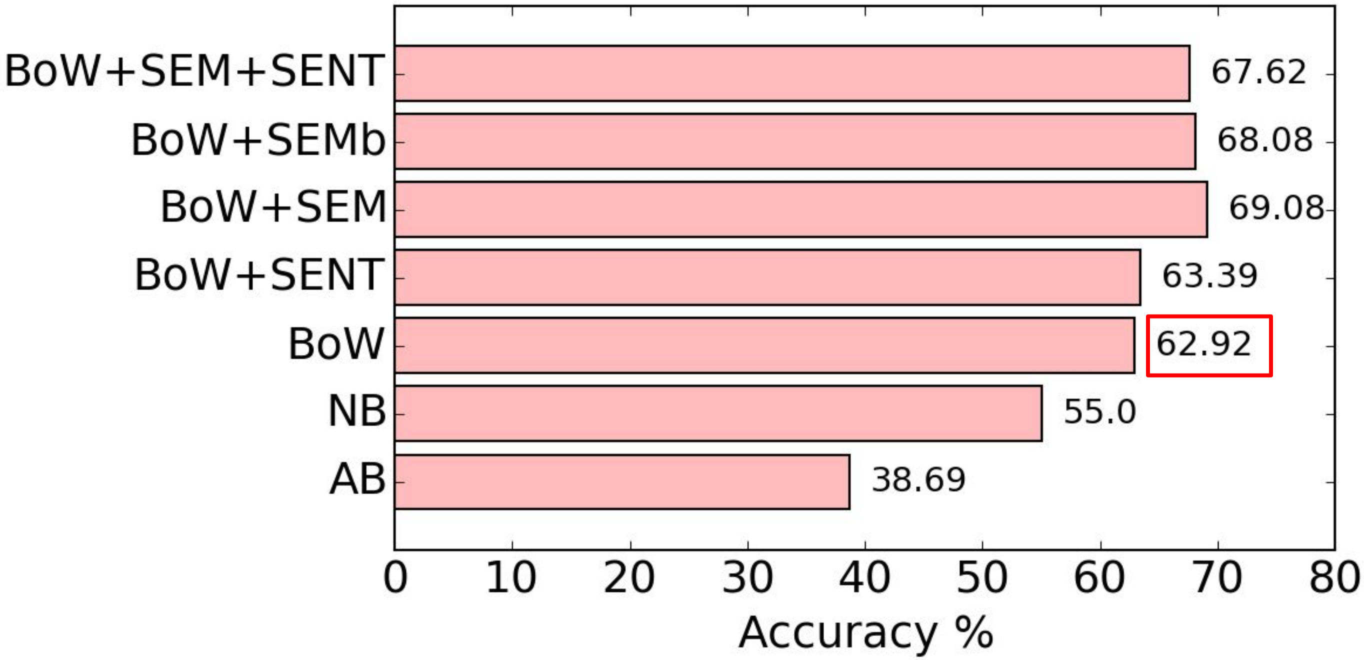
Genre Classification



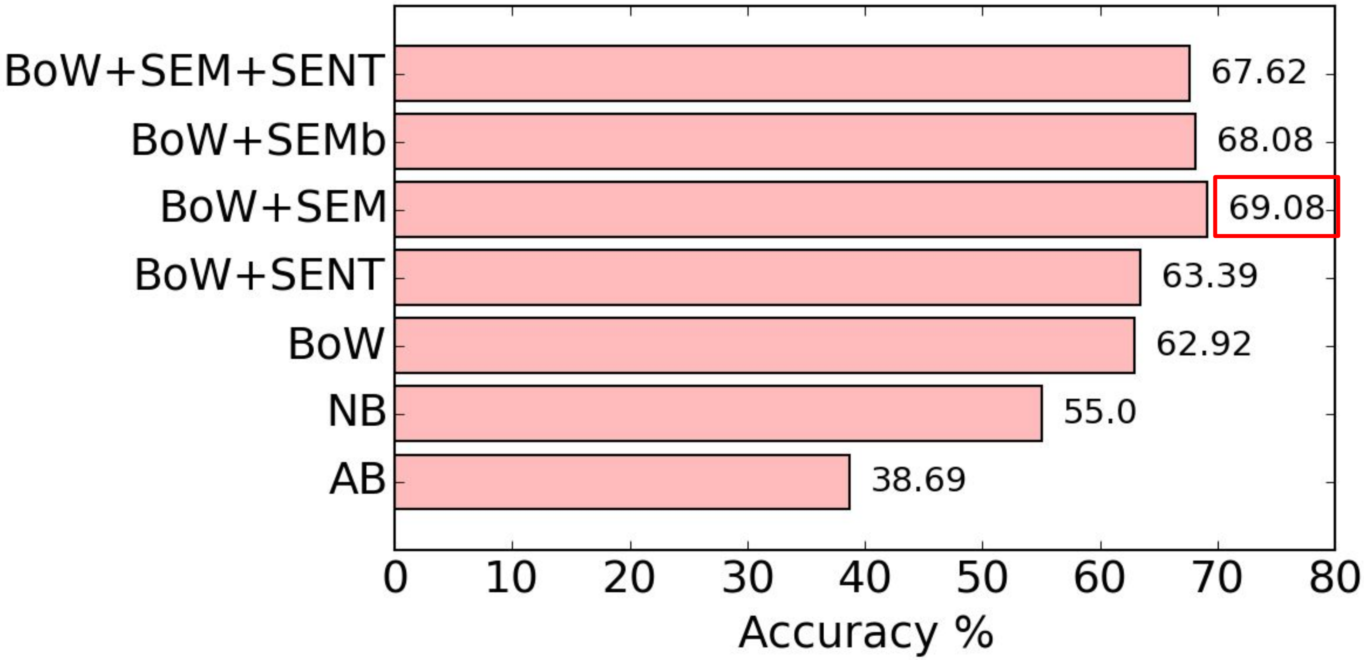
Genre Classification



Genre Classification



Genre Classification

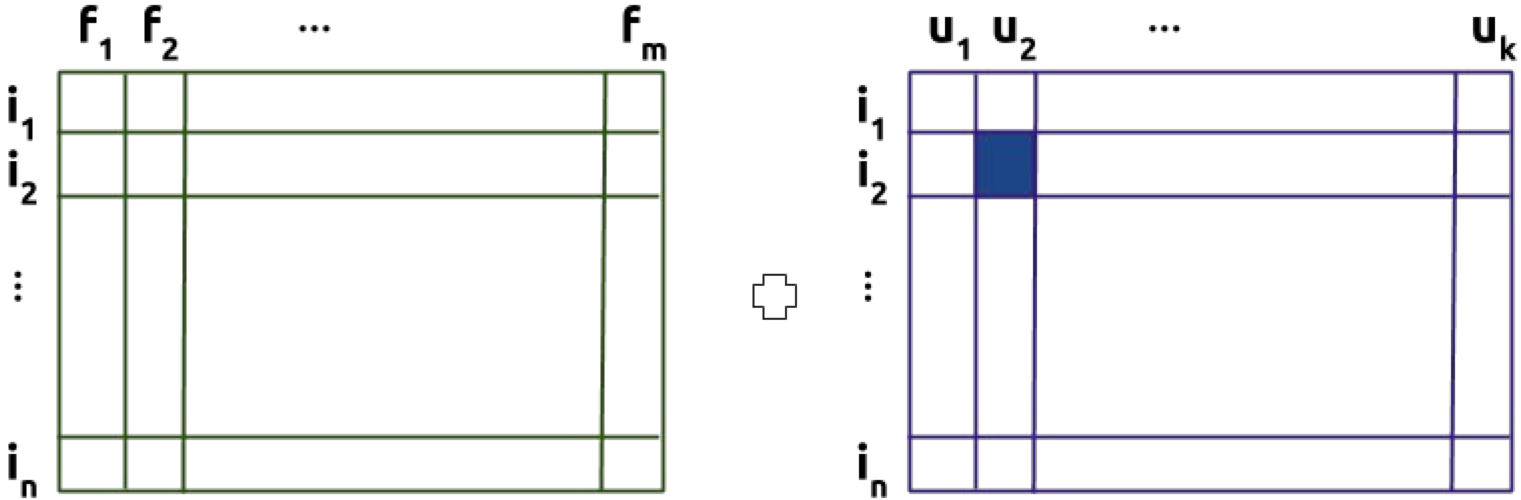


Genre Classification

Audio / Text

	Alt. Rock	Classical	Country	Electronic	Folk	Jazz	Latin	Metal	New Age	Pop	R&B	Hip-Hop	Rock
Alt. Rock	28 / 42	1 / 3	3 / 1	10 / 10	7 / 1	1 / 2	2 / 0	18 / 12	10 / 2	4 / 10	3 / 6	3 / 2	10 / 9
Classical	0 / 0	87 / 95	1 / 0	0 / 0	1 / 1	1 / 1	2 / 2	1 / 0	5 / 1	1 / 0	0 / 0	0 / 0	1 / 0
Country	2 / 1	0 / 0	51 / 84	3 / 0	9 / 1	9 / 0	3 / 0	0 / 1	3 / 0	8 / 8	6 / 4	1 / 0	5 / 1
Electronic	7 / 3	3 / 1	1 / 2	40 / 61	4 / 1	1 / 2	2 / 2	6 / 0	7 / 5	6 / 5	6 / 7	13 / 5	4 / 7
Folk	4 / 6	11 / 0	13 / 10	7 / 0	27 / 55	6 / 1	7 / 3	4 / 2	6 / 9	5 / 9	6 / 4	1 / 0	3 / 1
Jazz	7 / 0	10 / 1	6 / 2	2 / 2	5 / 0	45 / 82	6 / 3	3 / 0	8 / 2	3 / 5	4 / 1	1 / 1	0 / 1
Latin	4 / 3	6 / 4	9 / 2	1 / 2	5 / 1	10 / 2	28 / 78	3 / 0	6 / 2	11 / 4	7 / 2	5 / 0	5 / 0
Metal	13 / 5	1 / 0	1 / 1	2 / 2	1 / 0	0 / 1	1 / 0	63 / 87	1 / 0	1 / 0	3 / 1	1 / 0	12 / 3
New Age	9 / 2	7 / 6	9 / 0	7 / 4	10 / 10	9 / 2	7 / 6	3 / 3	15 / 53	10 / 7	6 / 1	2 / 1	6 / 5
Pop	6 / 2	9 / 1	10 / 2	9 / 2	5 / 3	9 / 2	5 / 2	2 / 0	7 / 1	19 / 73	7 / 6	2 / 2	10 / 5
R&B	8 / 2	0 / 1	16 / 3	8 / 4	2 / 0	5 / 3	5 / 0	1 / 0	3 / 0	7 / 10	24 / 71	17 / 5	4 / 1
Hip-Hop	8 / 2	0 / 0	2 / 1	8 / 2	0 / 1	0 / 1	1 / 0	4 / 3	2 / 0	4 / 1	7 / 2	61 / 86	3 / 1
Rock	17 / 15	1 / 2	6 / 8	4 / 7	10 / 5	2 / 4	7 / 1	12 / 13	4 / 1	9 / 7	7 / 4	6 / 2	15 / 31

Music Recommendation



Music Recommendation

Recommendation approaches:

- Collaborative filtering - only users matrix
- Content-based - only item features matrix
- Hybrid - both matrices

Music Recommendation

Hybrid approach: 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

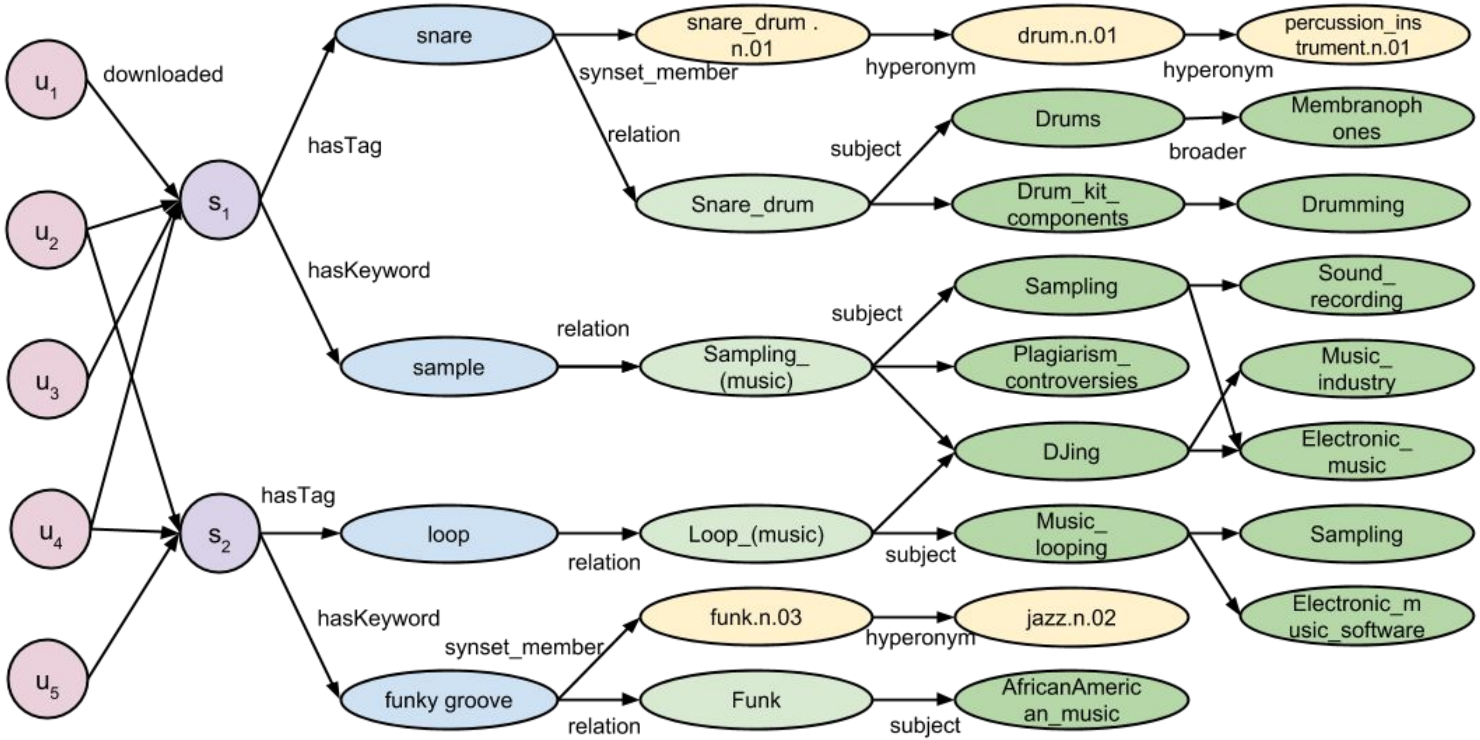
Music Recommendation

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 embedding: better novelty and diversity, slightly lower accuracy

Entity-based embedding: 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.

SONG #18

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

RECOMMENDED SONG

Fourth Time Around (Bob Dylan)

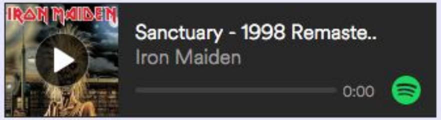
You Know My Name (Look Up The Number) <-- The Beatles <-- Fourth Time Around

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)

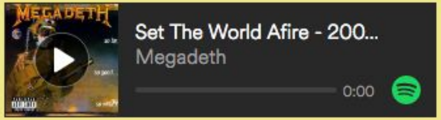


RECOMMENDED SONG

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

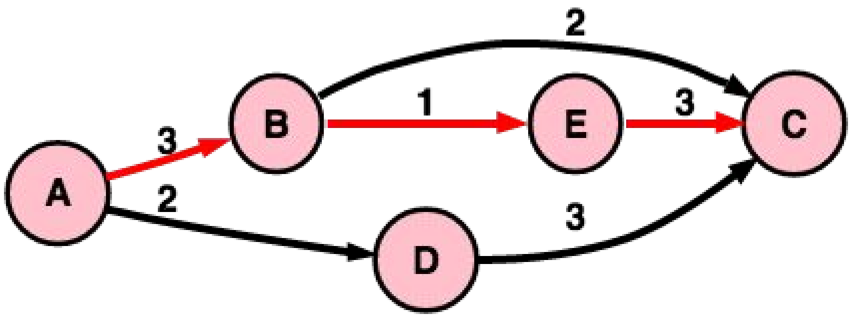


Give a score to the provided recommendation:
○ 1 ○ 2 ○ 3 ○ 4 ○ 5
Did you know the recommended song?
○ Yes ○ No

Interpreting Music Recommendations

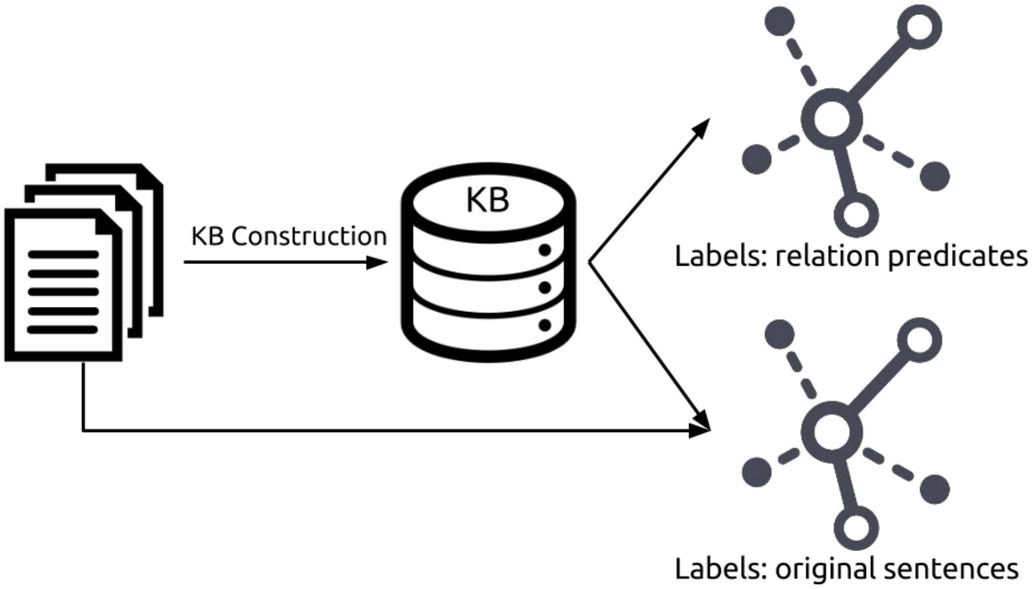
Challenges

- 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

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.

Sutcliffe, R. F. E., Crawford, T., Fox, C., Root, D. L., Hovy, E., & Lewis, R. (2015). Relating Natural Language Text to Musical Passages. Proceedings of the 16th International Society for Music Information Retrieval Conference, Malaga, Spain, 26-30 October, 2015

- 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-tutorial

Create a BabelNet account:

http://babelnet.org/register

Outline

- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - Semantic Enrichment of Musical Texts
- Applications in MIR
- **Applications in Musicology**
- Lexical Semantics
- Deep Learning
- Conclusions and Future

Applications in Musicology

Musicology

Musicology embraces the study of history, theory, and practice of music from many points of view.

Musicology is part of the humanities

Musicologists have to read a lot!

Musicology

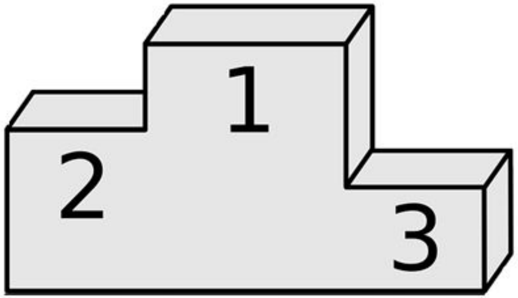
Musicology embraces the study of history, theory, and practice of music from many points of view.

Musicology is part of the humanities

Musicologists have to read a lot!

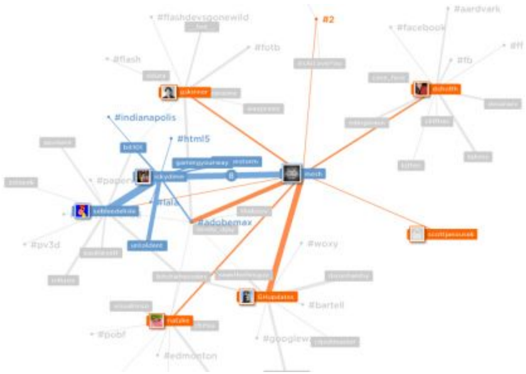
Applications in Musicology

Entity Relevance



Analytics

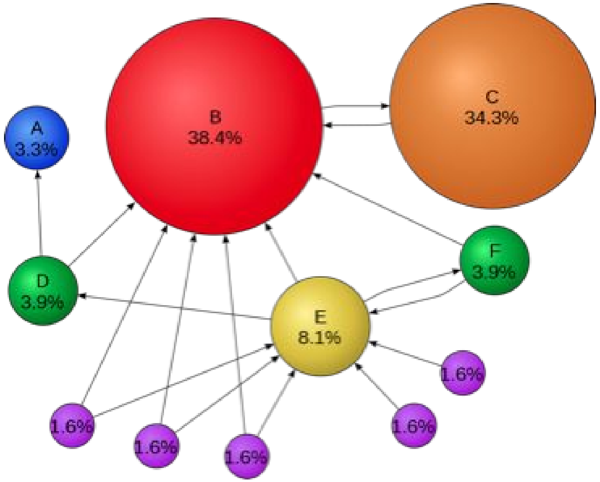
Information Visualization



Entity Relevance

See a Graph of Entities as network of hyperlinks

Use Pagerank or HITS to compute entity relevance



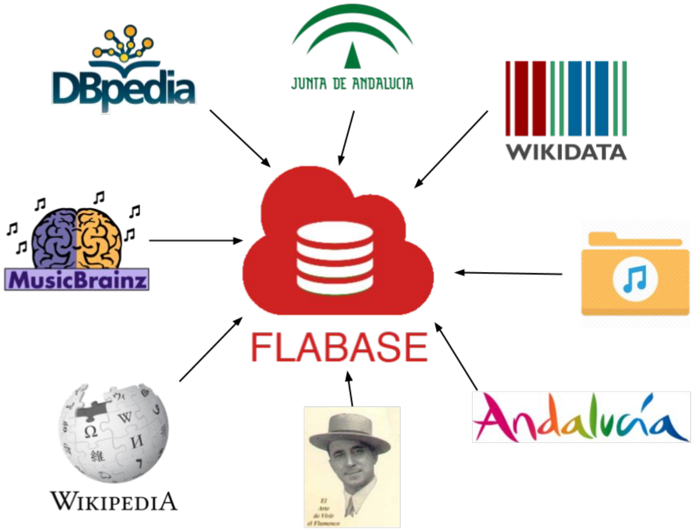
Wilco
This alternative rock band was formed in 1994 by the remaining members of Uncle Tupelo following singer Jay Farrar's departure.

FlaBase

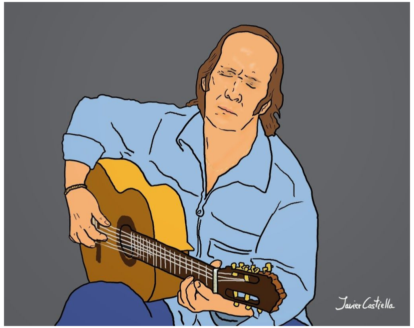
Oramas S., Gómez F., Gómez E., & Mora J. (2015). FlaBase: Towards the creation of a Flamenco Music Knowledge Base. 16th International Society for Music Information Retrieval Conference (ISMIR 2015).

1,174 Artists (text biography)
76 Palos (flamenco genres)
2,913 Albums
14,078 Tracks
771 Andalusian locations

We built a **Graph of Entities**



Artist Relevance



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

Analytics

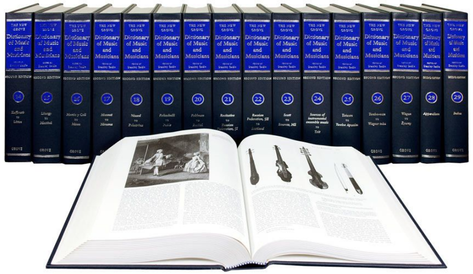
- Extract attributes, events, entity mentions, relations, sentiment, etc.
- Compute analytics

Useful insights for musicologists

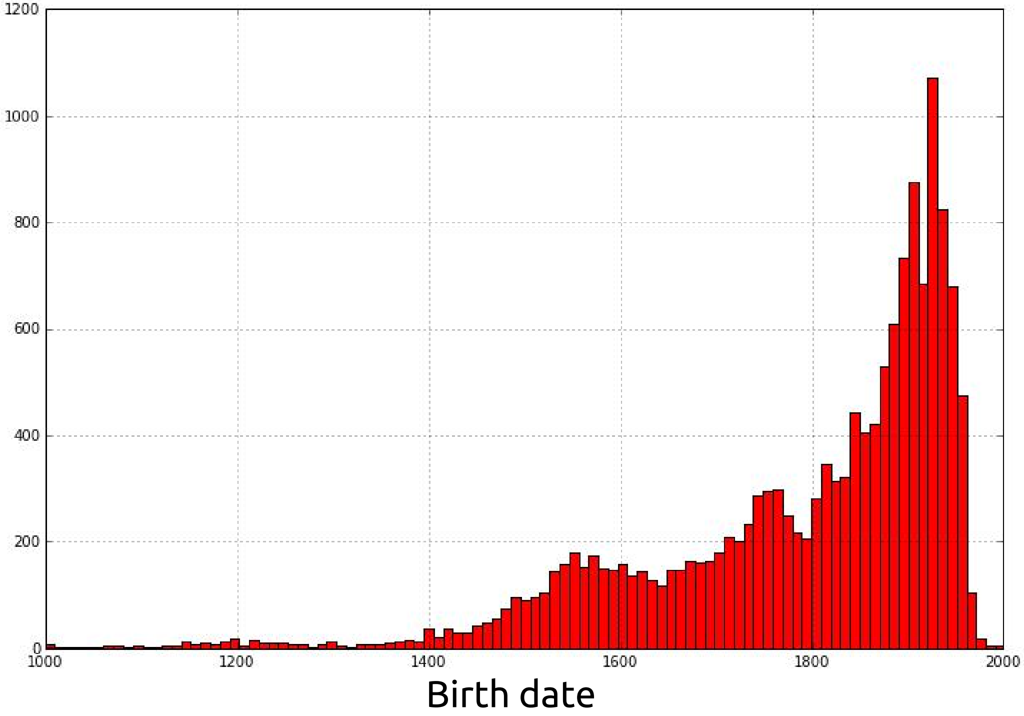
Analytics: Grove Dictionary

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.

- **Grove Dictionary**: one of the largest reference works on Western music
- 16,707 biographies were gathered from Grove Music Online
- **Extracted information**: roles, birth and death, entity mentions, relations



Analytics: Grove Dictionary



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

Analytics: Grove Dictionary

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: Grove Dictionary

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: Diachronic study of affective language

MARD Multimodal Album Reviews Dataset

- Amazon (~66k albums / ~250k customer reviews)
  - Album customer reviews
  - Genre tags (16 genres and 287 subgenres)
  - Star Ratings
  - Metadata: title, artist, record label
- MusicBrainz: ids, song titles, year of publication
- AcousticBrainz: audio descriptors of songs

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

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/)

Diachronic Study of Affective Language

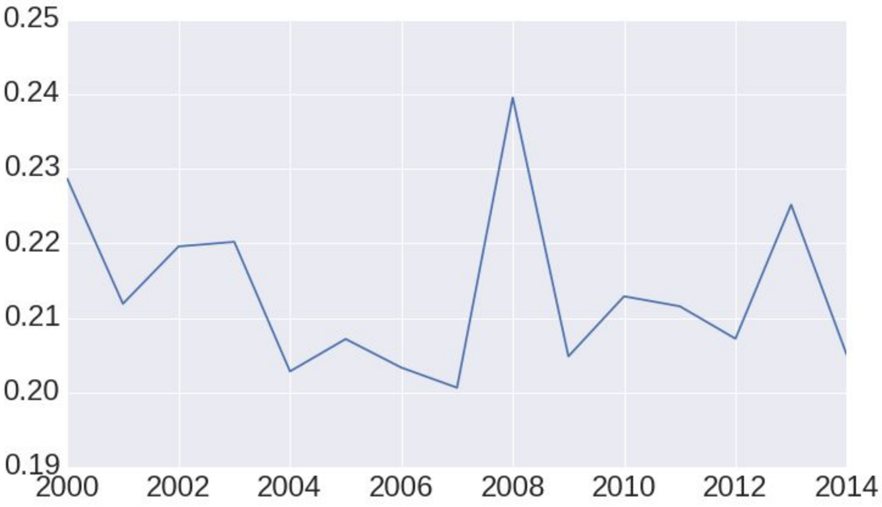
Sentiment score: Average sentiment score of all aspects in a review

Two perspectives:

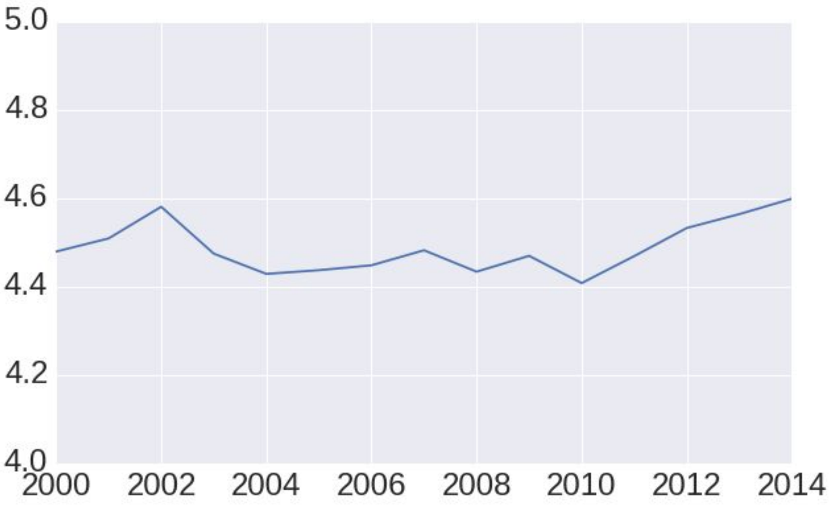
- Average of all reviews by review publication year (2000-2014)
  - Evolution of affective language from a customer perspective
- Average of all reviews by album publication year (1950-2014)
  - Evolution of affective language from a musical perspective

Study by **review** publication year

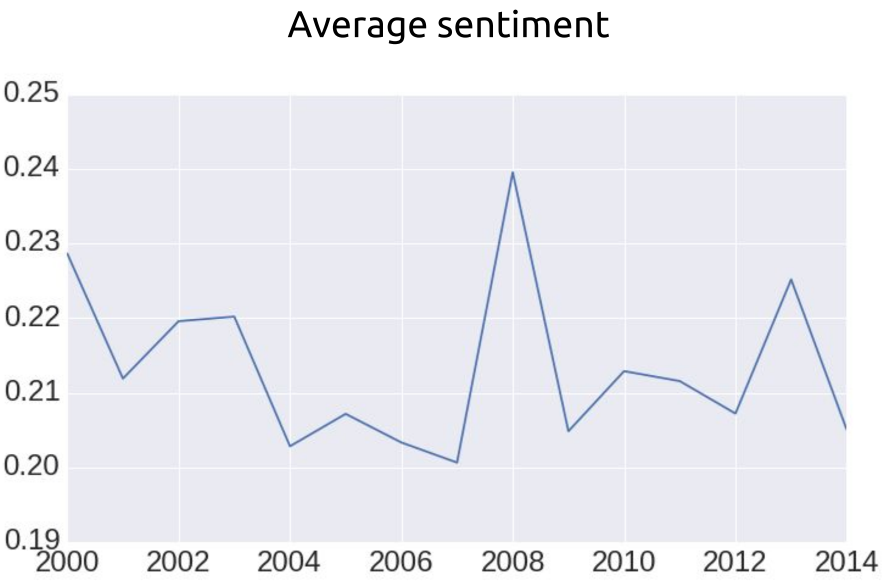
Average sentiment

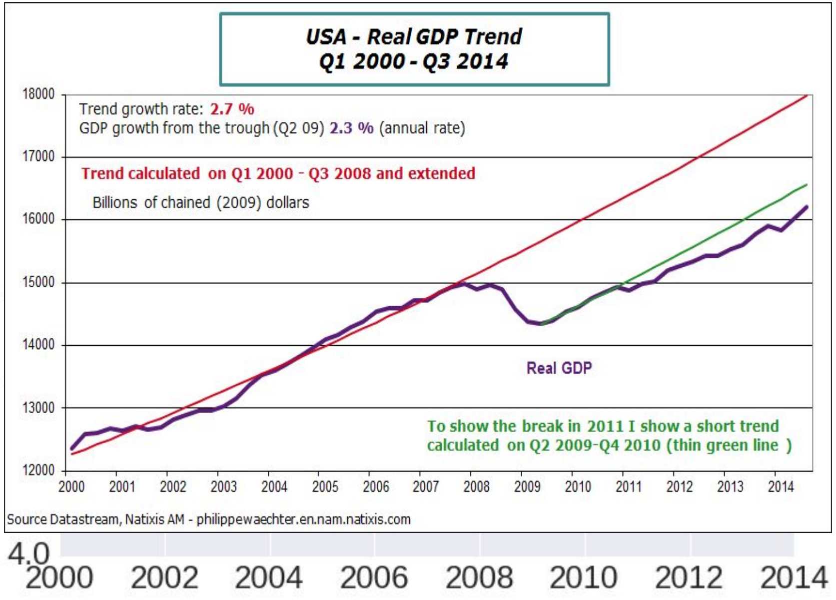


Average rating



Study by **review** publication year





Study by **review** publication year

Dominique Moïsi in:

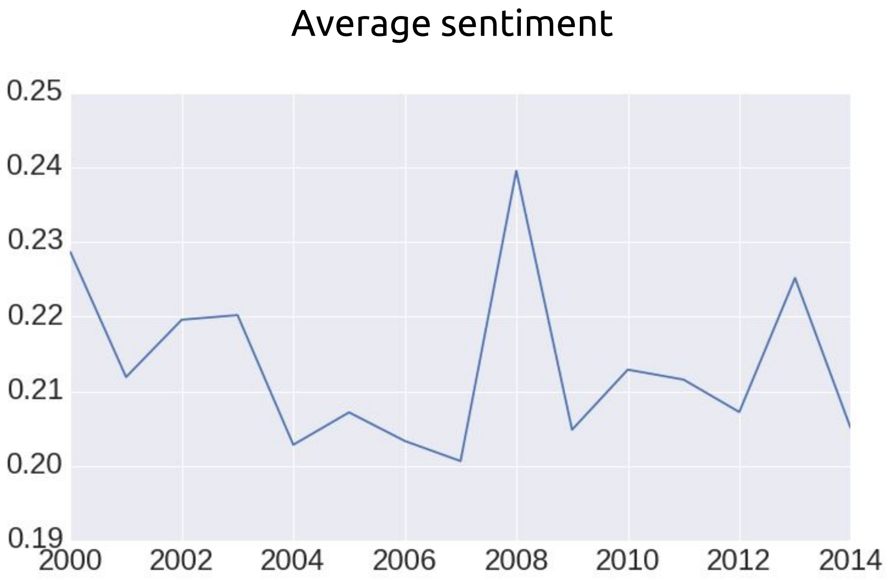
In November 2008, at least for a time, hope prevailed over fear. The wall of racial prejudice fell as surely as the wall of oppression had fallen in Berlin twenty years earlier [...] Yet the emotional dimension of this election and the sense of pride it created in many Americans must not be underestimated.

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:

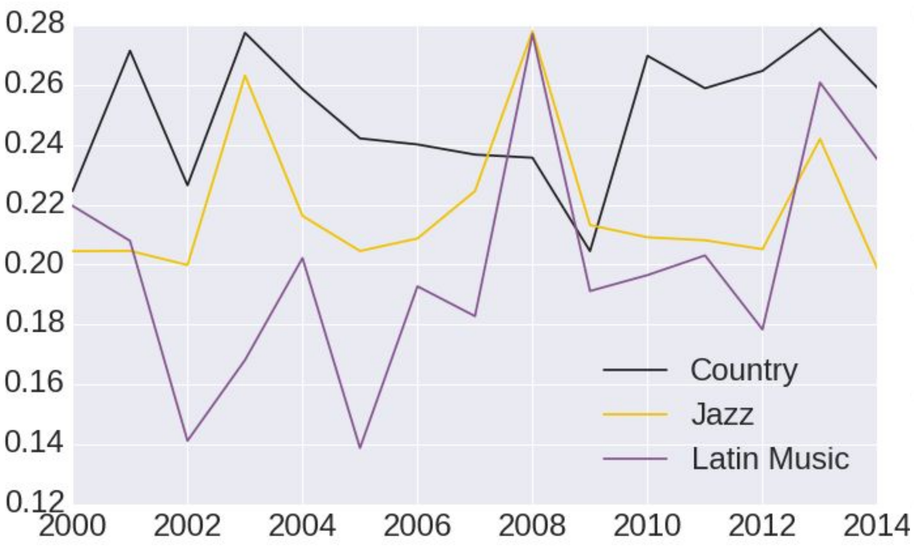
In November 2008, at lea … he wall of racial prejudice fell as surely … rlin twenty years earlier [...] Yet the emo … sense of pride it created in many Ame



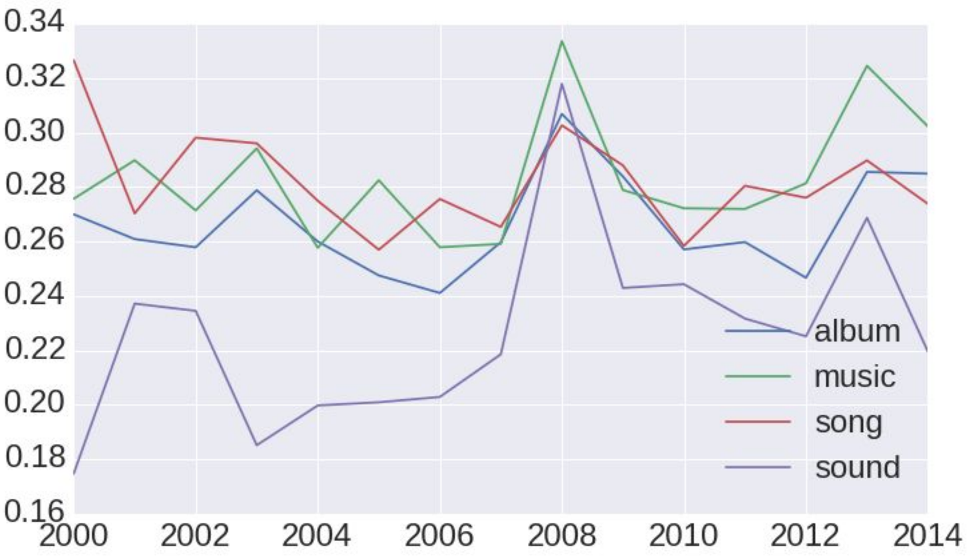
Dominique Moisi. The Geopolitics of Emotic … g the World. Anchor Books, New York, NY, USA, 2010.

Study by **review** publication year

Average sentiment by genre



Average sentiment by aspect



Study by **review** publication year

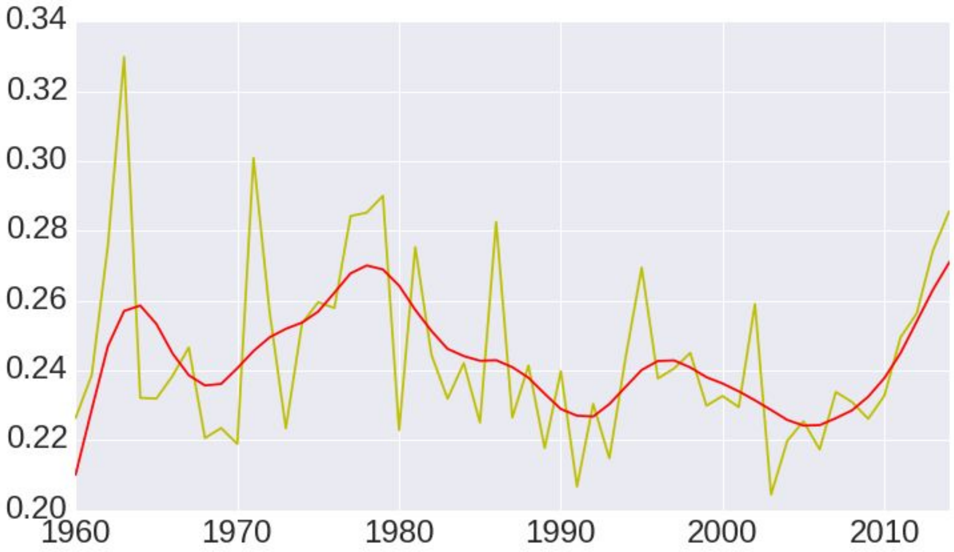
Further studies necessary to validate any of these suggestions

Correlation ≠ Causation

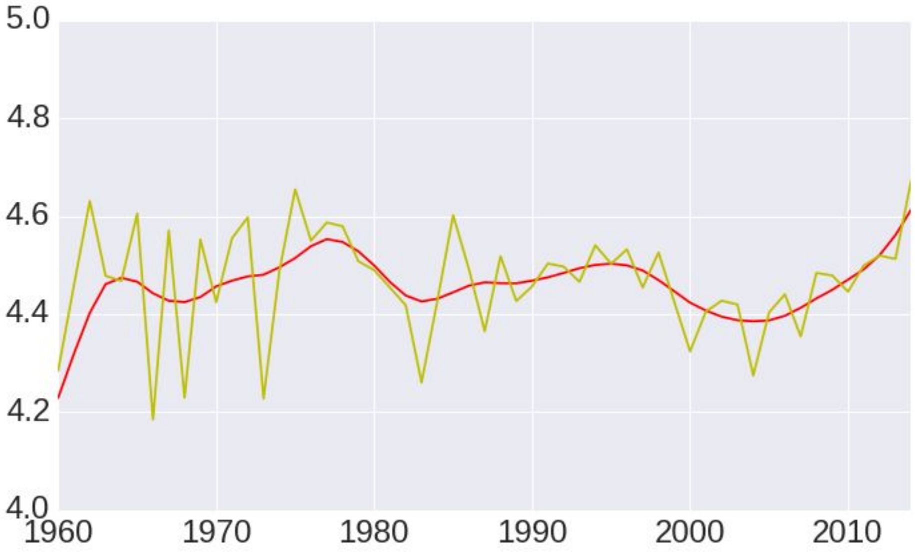
Interesting insight for Musicologists

Study by **album** publication year

Average sentiment



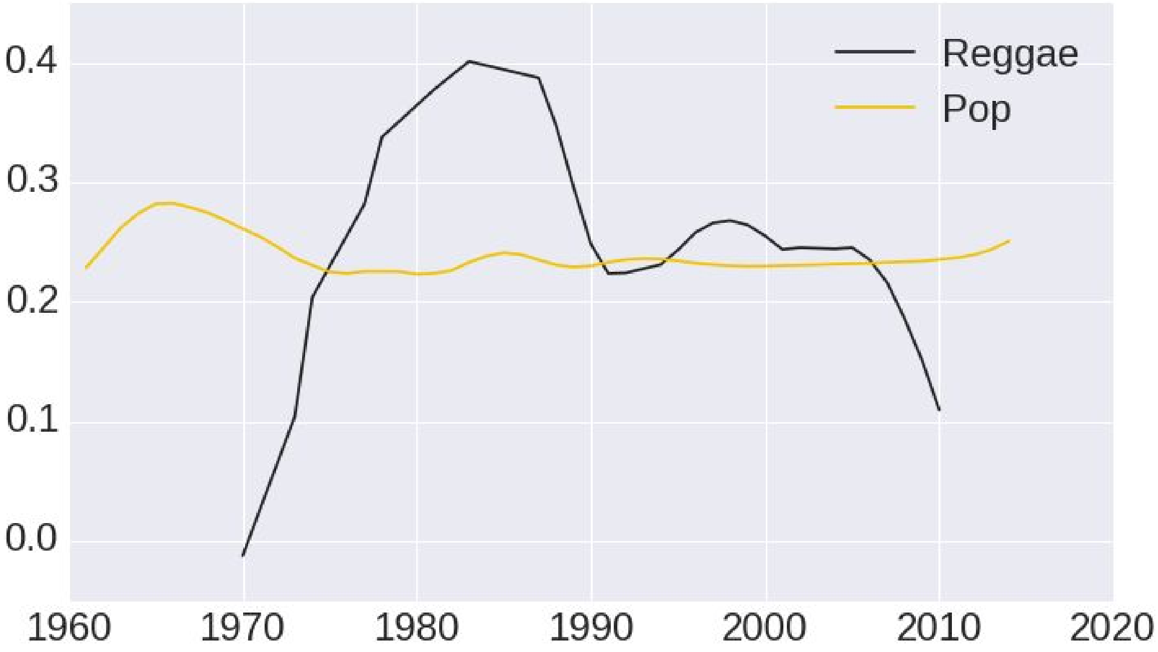
Average rating



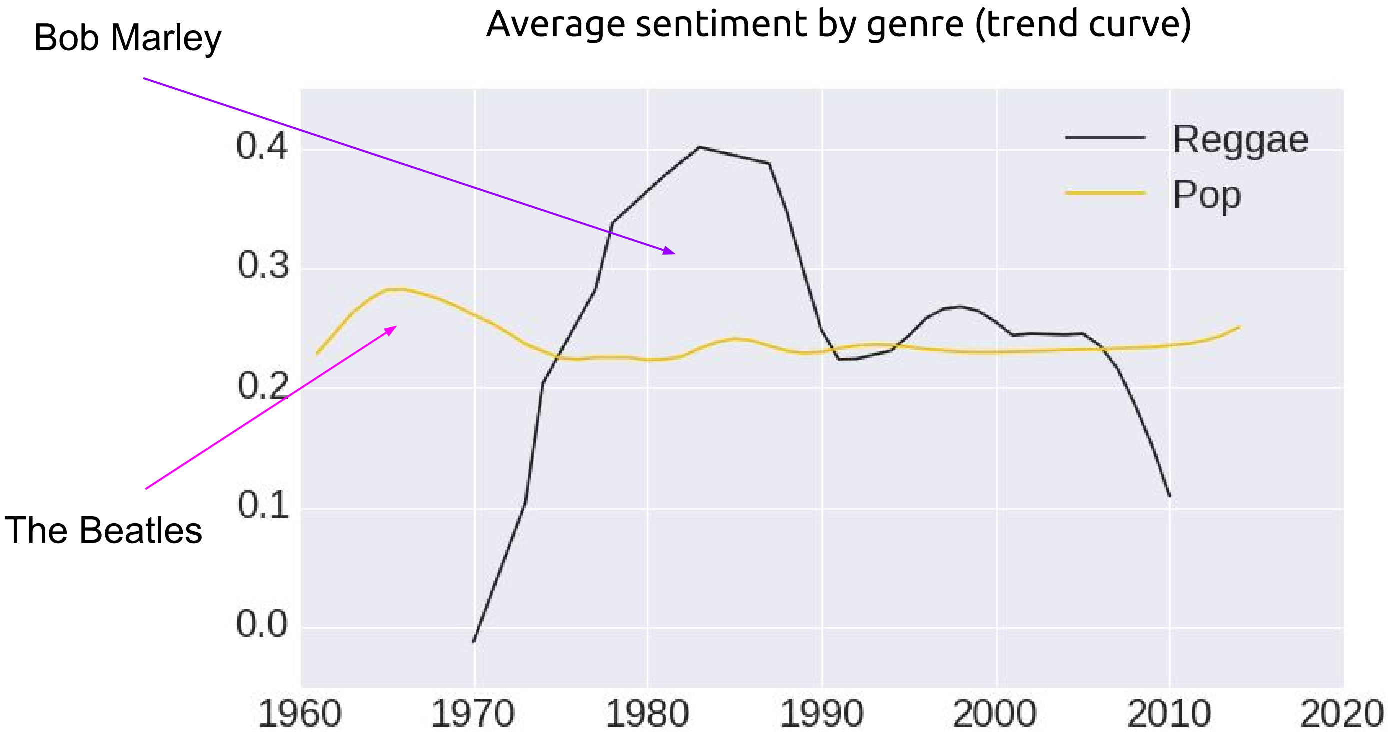
Pearson's correlation r = 0.75, $p \ll 0.001$

Study by **album** publication year

Average sentiment by genre (trend curve)



Study by **album** publication year



Study by **album** publication year

Approach useful to study evolution of music genres

Strong correlation between average sentiment and average rating

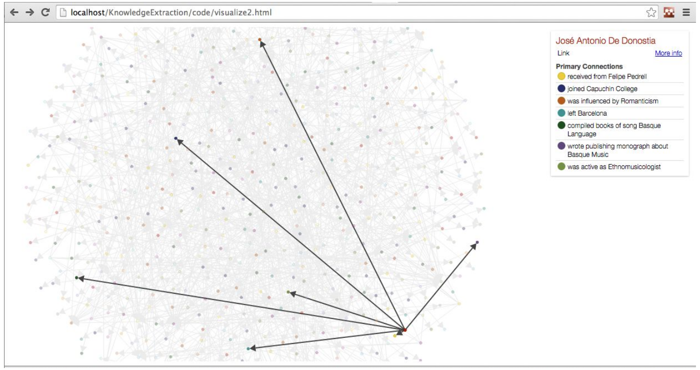
Again useful insights for musicologists

Information Visualization

Extract a Knowledge Base from the documents of 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)

Outline

- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- **Lexical Semantics**
- Deep Learning
- Conclusions and Future

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: formal (logic), path-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).

wampimuk

Distributional Lexical Semantics

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

He filled the **wampimuk** with the substance, passed it around we all drunk some.

Distributional Lexical Semantics

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

He filled the **wampimuk** with the substance, passed it around we all drunk some.

We found a little, hairy **wampimuk** sleeping behind the tree.

Distributional Lexical Semantics

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

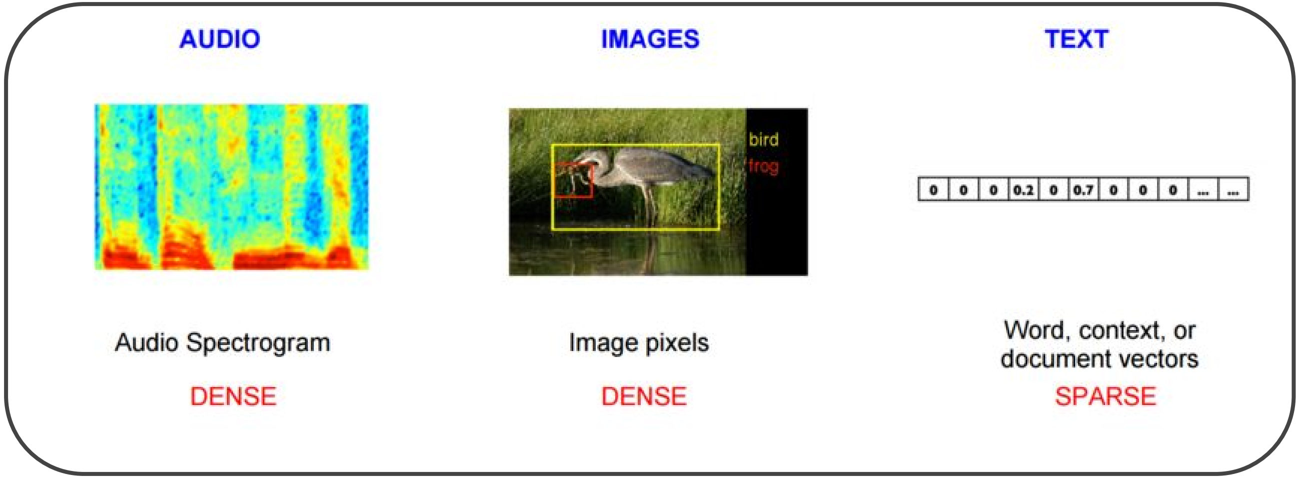
He filled the **wampimuk** with the substance, passed it around we all drunk some.

We found a little, hairy **wampimuk** sleeping behind the tree.

(McDonald and Ramscar, 2001)

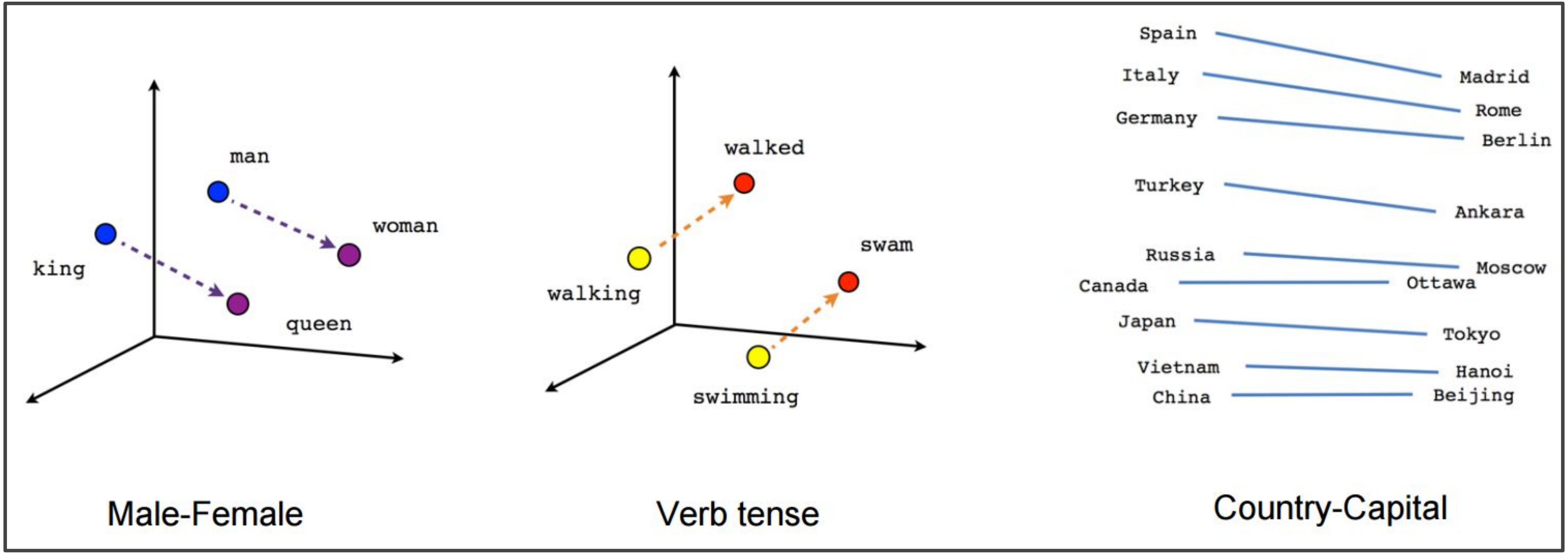
Distributional Hypothesis: words that appear in similar contexts exhibit similar semantics.

Distributional Lexical Semantics



- Project linguistic items in vector space.
- Predictive models vs count-based models (Baroni et al., 2014).
- **word2vec** (Mikolov et al., 2013), Glove (Pennington et al., 2014) …

Distributional Lexical Semantics



Distributional Lexical Semantics

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

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

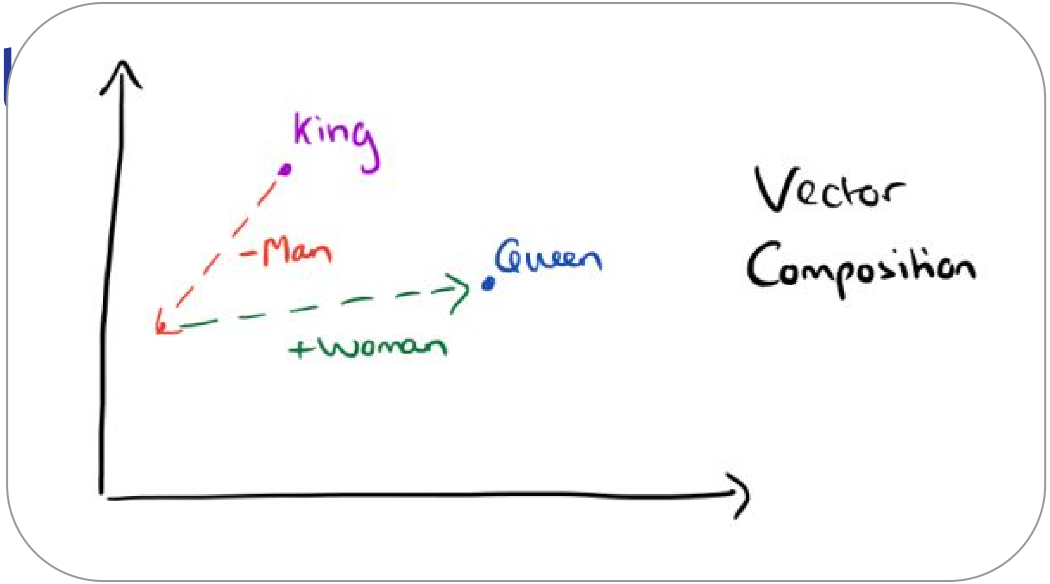
gensim

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

Distributional Lexical

· Word similarity, relatedness

or **analogy** tasks.




```
>>> 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)]
```

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) …]

Other uses of embeddings for music lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.

Other uses of embeddings for lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.
  - The influence of *sisters of mercy* became evident in later *poetry*.

Other uses of embeddings for lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.

  - The influence of ***sisters of mercy*** became evident in later *poetry*.

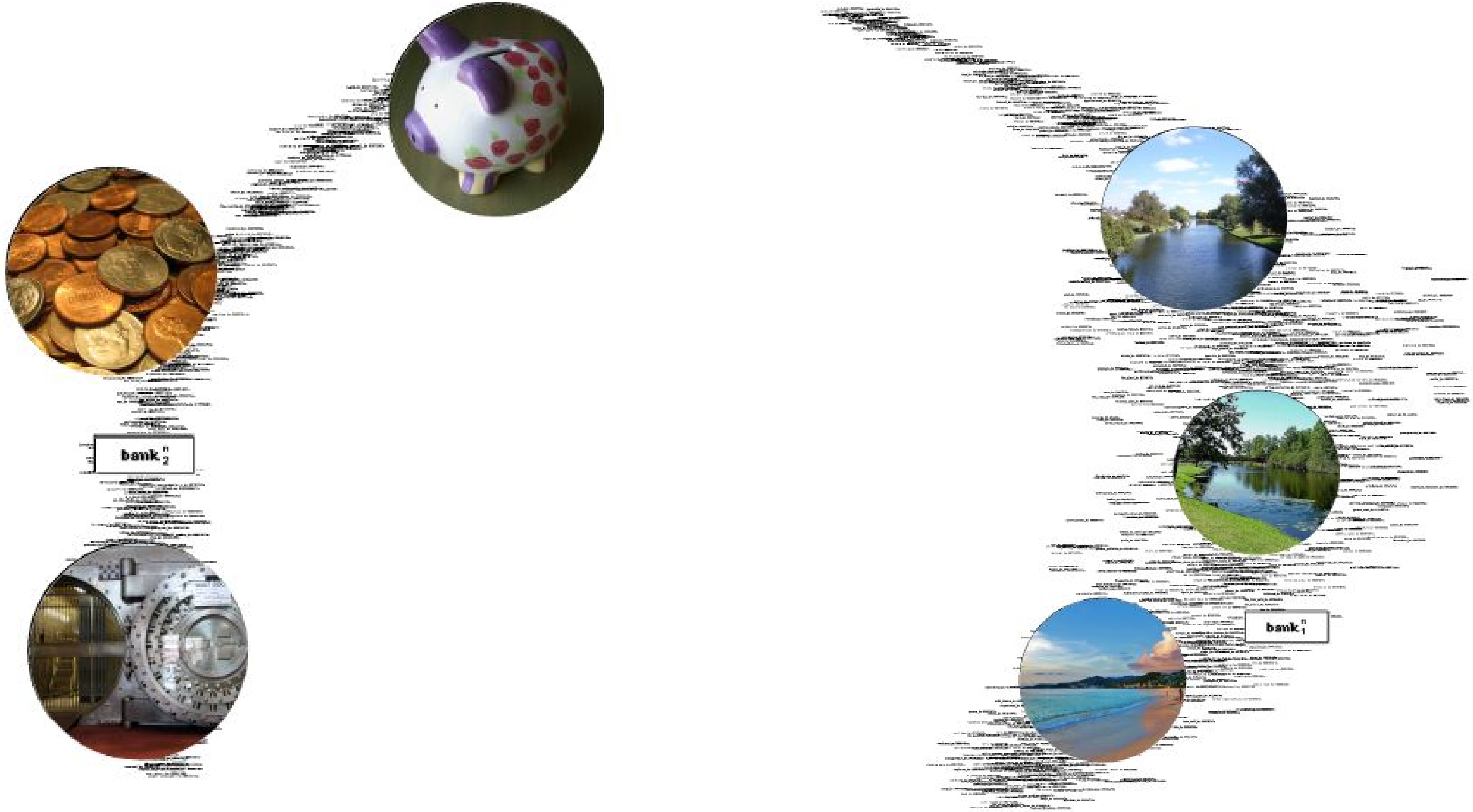
Other uses of embeddings for lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.
    - The influence of *sisters of mercy* became evident in later *poetry*.

Exploit sense-level embeddings using **BabelNet** (Navigli and Ponzetto, 2012) as a reference sense inventory (e.g. SensEmbed, by Iacobacci et al. 2015)

https://iiacobac.wordpress.com/2015/09/02/sensembed/

BabelNet



$bank^{n}_{2}$
$bank^{n}_{1}$

Other uses of embeddings for lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.
  - The influence of *sisters of mercy* is evident in many later *poetry* acts.

"sisters of mercy"

(Delli Bovi et al., 2015)

$$\cos(n, \varphi) = \max_{v_n \in V(n), v_\varphi \in V(\varphi)} \frac{v_n \cdot v_\varphi}{||v_n|| \, ||v_\varphi||}$$

"poetry"

{sisters_of_mercy$_{bn1}$, sisters_of_mergy_(song)$_{bn2}$...}

{poetry$_{bn1}$ ...}

Other uses of embeddings for lexical semantics

- Word Sense Disambiguation and Entity Linking in the music domain.

```
>>> import sensembed_api as sensembed
>>> sister_senses = sensembed.getLemmaSenses('sisters_of_mercy')
>>> sisters_senses
[u'sisters_of_mercy_bn:00424887n', u'sisters_of_mercy_bn:03828439n']
>>> poetry_senses = sensembed.getLemmaSenses('poetry')
>>> sensembed.closest_senses(sisters_senses, poetry_senses)
(u'sisters_of_mercy_bn:03828439n', u'poetry_bn:00063195n',
0.0894221650994795 2)
```

Other uses of embeddings for lexical semantics

· Word Sense Disambiguation and Entity Linking in the music domain.

```
>>> import sensembed_api as sensembed
>>> sister_senses = sensembed.getLemmaSenses('sisters_of_mercy')
>>> sisters_senses
[u'sisters_of_mercy_bn:00424887n', u'sisters_of_mercy_bn:03828439n']
>>> poetry = sensembed.getLemmaSenses('poetry')
>>> sensembed.closest_senses(sisters_senses, poetry)
(u'sisters_of_mercy_bn:03828439n', u'poetry_bn:00063195n', 0.08942216509947952)
```

Conclusion

· Lexical semantics is a *buzzword* in NLP.

· VSMs, lexical semantics and advances in neural approaches have opened up a vibrant area of research.

· EMNLP2015 (Conference with A rating according to Google Scholar):

   * **Empirical Methods in Natural Language Processing**

   * "The insider joke in Lisbon was that the E in EMNLP now stands for Embedding (instead of Empirical) (…) " (https://wit.ai/blog/2015/09/23/emnlp)

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).

Outline

- Introduction to NLP
- Information Extraction
  - Construction of Music Knowledge Bases
  - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- **Deep Learning**
- Conclusions and Future

Deep Learning

Deep Learning in Natural Language Processing

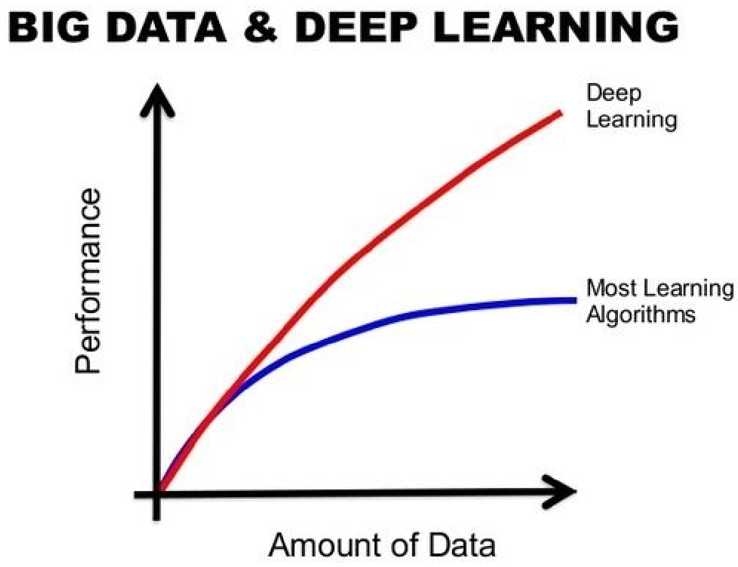
Deep Learning improves almost all tasks in NLP!! (as in many other fields)

Deep Network Architectures: LSTM y CNN

LSTM: parsing, entity recognition, sentiment analysis

CNN: classification, sentiment analysis

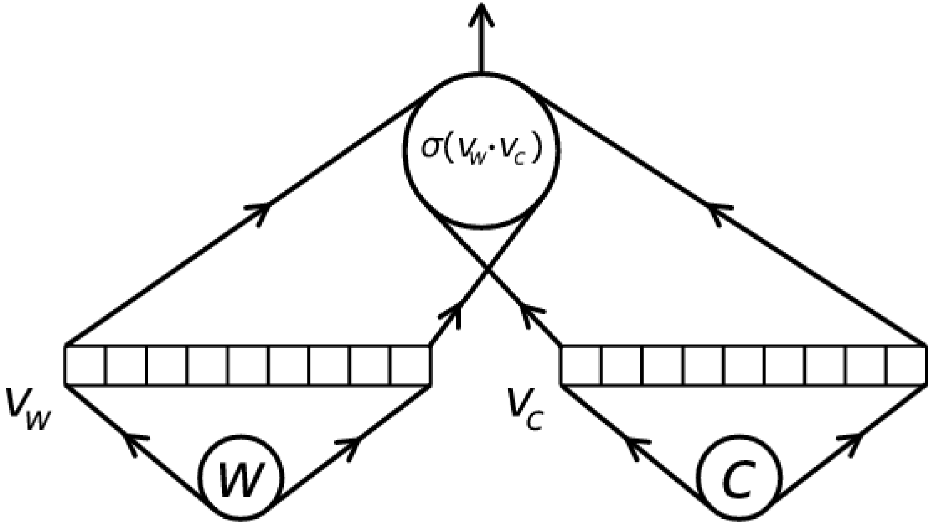
More than words: end-to-end, character level processing, word embeddings



Word2vec

Predict a context word $\mathbf{c} \in (\mathbf{w}_{i-L}, \ldots, \mathbf{w}_{i-1}, \mathbf{w}_{i+1}, \ldots, \mathbf{w}_{i+L})$ 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

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

According to Levy et al. 2014

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

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

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.

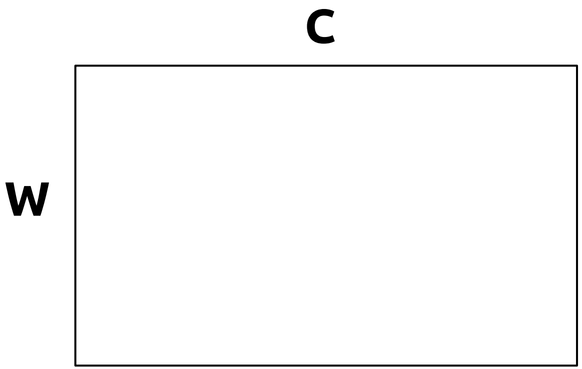
O. Levy, Y. Goldberg (2014): Neural Word Embedding as Implicit Matrix Factorization. NIPS 2014

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

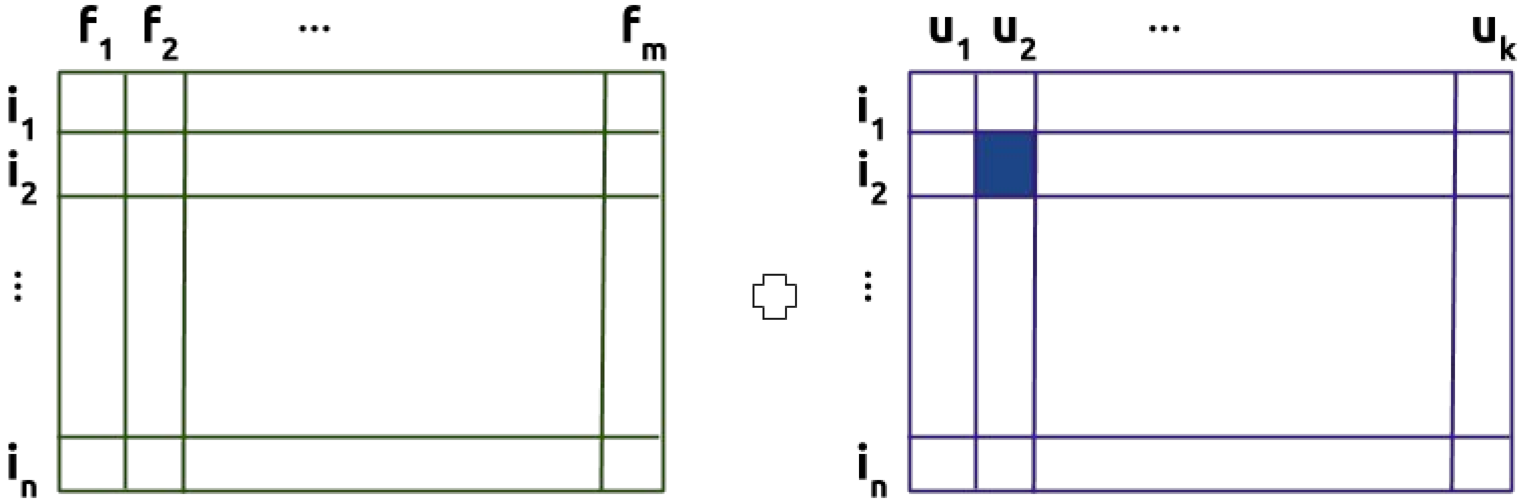
Trained with Gensim in Art of the Mix playlists
(http://labrosa.ee.columbia.edu/projects/musicsim/aotm.htm)

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)]

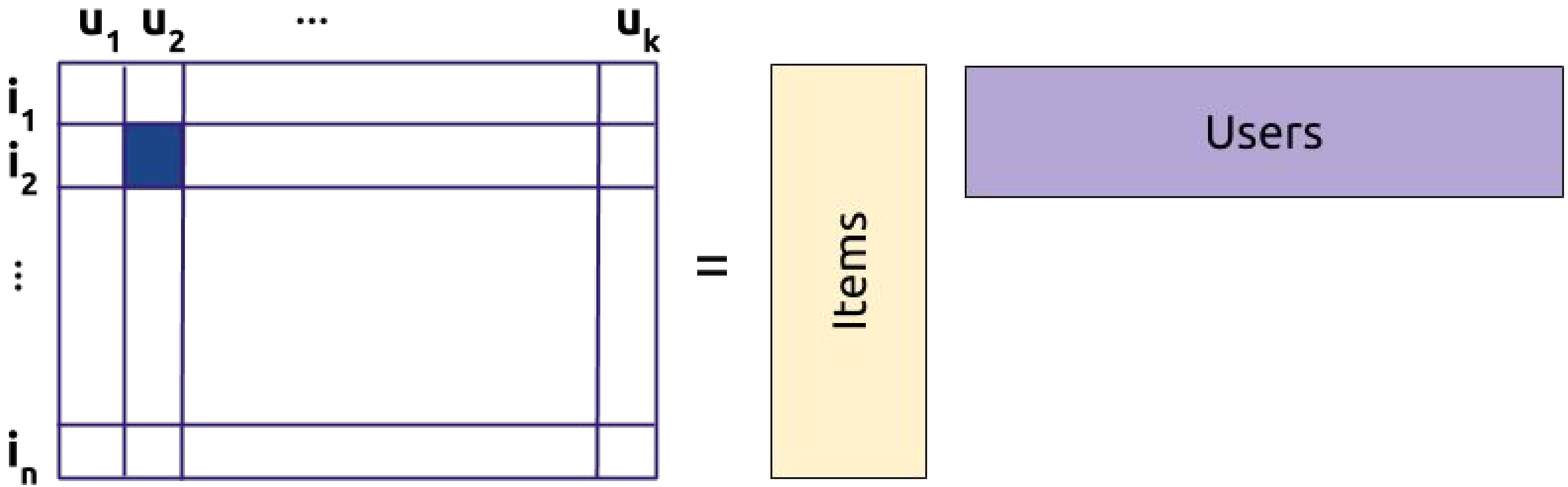
model.most_similar('nirvana')
[('soundgarden', 0.84231329), ('pearl jame', 0.8271907), ('oysterhead', 0.81855756)]

Deep Learning for Music Recommendation



Deep Learning for Music Recommendation

Matrix Factorization

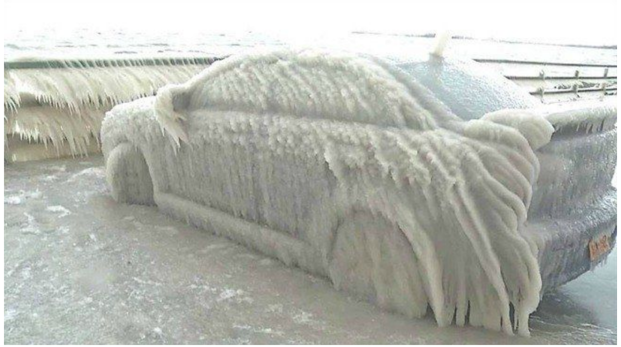


Cold start problem

No user's information for new items → Collaborative filtering doesn't work

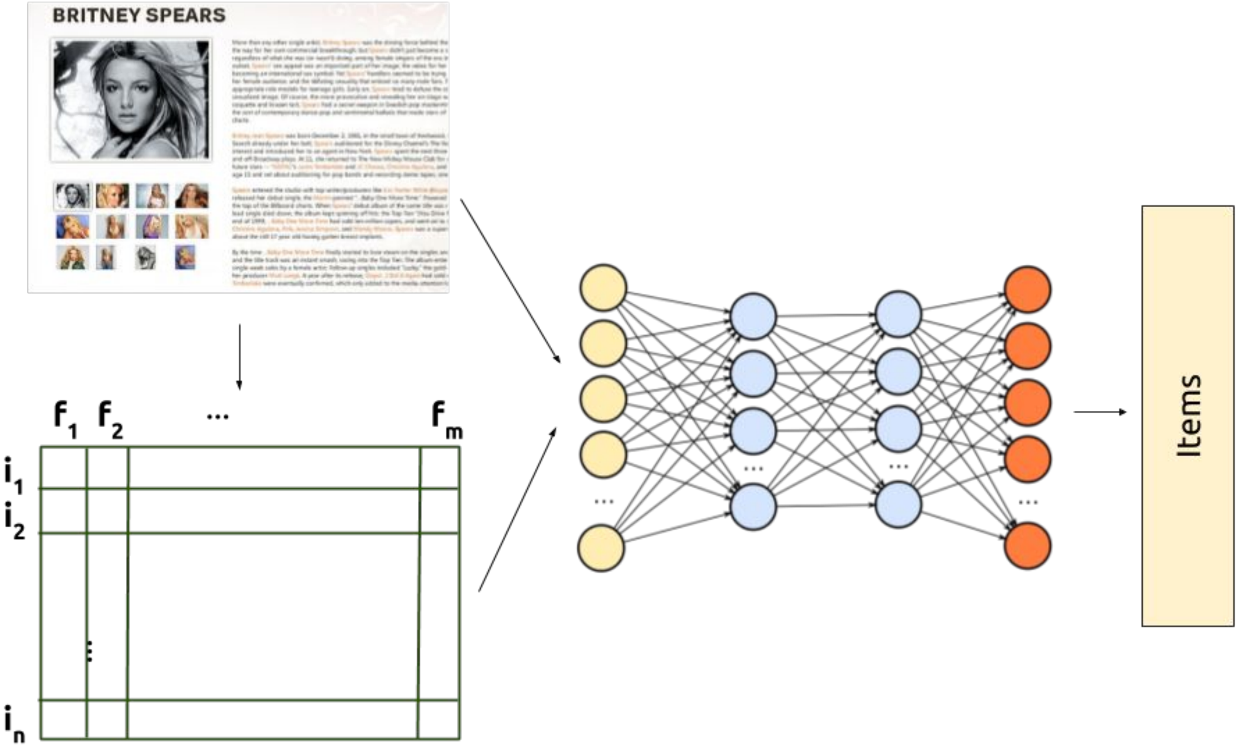
Need of content-based or hybrid approaches:

- Aggregation of feature vectors
- **Learn item factors from content features**



Aäron van den Oord, Sander Dieleman, and Benjamin Schrauwen. 2013. Deep content-based music recommendation. In Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS'13)

Deep Learning for Music Recommendation



Dataset

Million Song Dataset + Artist biographies and tags from **Last.fm**

Artists: ~27k
Users: 1 million
Sparsity: 0.9990

Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. In Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011.

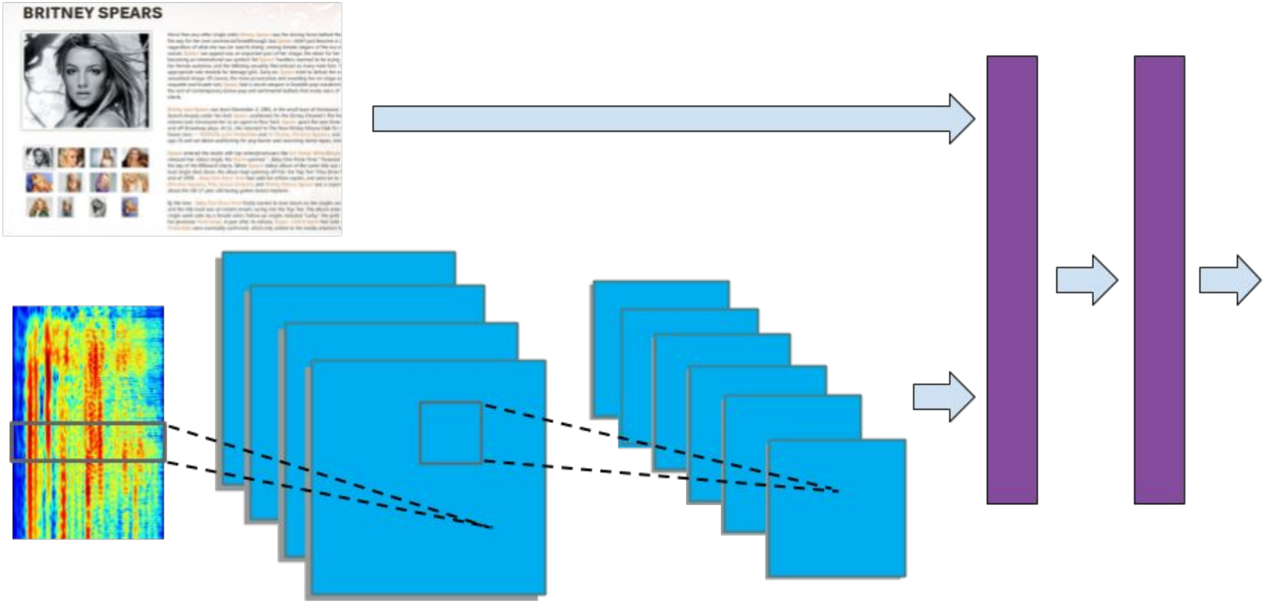
Input	Embed.	Learning	MAP@500	ROC-AUC
Text	VSM	Random Forest	0.015	0.664
Text	VSM	Feed Forward	0.030	0.748
Text + Semantic Graph	**VSM**	**Feed Forward**	**0.035**	**0.748**
Text	avg-w2v	Feed Forward	0.010	0.686
Text	w2v	LSTM	0.010	0.697
Text + Semantic Graph	n2v	Feed Forward	0.028	0.763
Random	-	-	0.001	0.495
Tags	VSM	Feed Forward	0.057	0.786
Upperbound	-	-	0.519	0.955

Input	Embed.	Learning	MAP@500	ROC-AUC
Text	VSM	Random Forest	0.015	0.664
Text	VSM	Feed Forward	0.030	0.748
Text + Semantic Graph	**VSM**	**Feed Forward**	**0.035**	**0.748**
Text	avg-w2v	Feed Forward	0.010	0.686
Text	w2v	LSTM	0.010	0.697
Text + Semantic Graph	n2v	Feed Forward	0.028	0.763
Random	-	-	0.001	0.495
Tags	VSM	Feed Forward	0.057	0.786
Upperbound	-	-	0.519	0.955

Input	Embed.	Learning	MAP@500	ROC-AUC
Text	VSM	Random Forest	0.015	0.664
Text	VSM	Feed Forward	0.030	0.748
Text + Semantic Graph	**VSM**	**Feed Forward**	**0.035**	**0.748**
Text	avg-w2v	Feed Forward	0.010	0.686
Text	w2v	LSTM	0.010	0.697
Text + Semantic Graph	n2v	Feed Forward	0.028	0.763
Random	-	-	0.001	0.495
Tags	VSM	Feed Forward	0.057	0.786
Upperbound	-	-	0.519	0.955

Multimodal Approach

Audio and text can be combined in a deep neural network



Outline

- Introduction to NLP
- Information Extraction
    - Construction of Music Knowledge Bases
    - Semantic Enrichment of Musical Texts
- Applications in MIR
- Applications in Musicology
- Lexical Semantics
- Deep Learning
- **Conclusions and Future**

Conclusions and Future

Conclusions

- The extraction of high level semantic representations from text have been shown useful in different MIR and Musicological problems.
- There is already a need of new methodologies that better exploit these semantic representations.
- Word Embeddings and Deep Learning opens a new world of barely exploited possibilities.
- This tutorial is an initial attempt to boost the interaction between the NLP and MIR communities.

Datasets Overview

Name	Documents	Task	Link
SAS	artist biographies	similarity	http://mtg.upf.edu/download/datasets/semantic-similarity
MARD	album reviews	classification	http://mtg.upf.edu/download/datasets/mard
KGRec-sound	sound descriptions	recommendation	http://mtg.upf.edu/download/datasets/knowledge-graph-rec
KGRec-music	song stories	recommendation	http://mtg.upf.edu/download/datasets/knowledge-graph-rec
ELMD	artist biographies	entity recognition	http://mtg.upf.edu/download/datasets/elmd

KBs Overview

Name	Source documents	Link
KBSF	songs stories	http://mtg.upf.edu/download/datasets/kbsf
FlaBase	flamenco music webs	http://mtg.upf.edu/download/datasets/flabase

Open Knowledge Extraction Challenge @ European Semantic Web Conference'17

· We are currently annotating and validating a gold standard dataset in the context of Task 3 in the OKE challenge @ ESWC 2017:

- Focused Musical NE Recognition and Linking

· A good opportunity to develop and evaluate an EL system in the music domain.

· Reference inventory is MusicBrainz (instead than the classic DBpedia URIs).

EXCELENCIA MARÍA DE MAEZTU

Open Knowledge Extraction Challenge @ European Semantic Web Conference’17

Call for Participation - 2 Tasks

Musical NE Recognition

Identification of musical entities: Artist, Album, Song

Musical NE Linking

Linking of identified entities to MusicBrainz

https://project-hobbit.eu/challenges/oke2017-challenge-eswc-2017/

Open Knowledge Extraction Challenge @ European Semantic Web Conference'17

Call for Participation - 2 Tasks

Musical NE Recognition

Identification of musical entities: Artist, Album, Song

Musical NE Linking

Linking of identified entities to MusicBrainz

https://project-hobbit.eu/challenges/oke2017-challenge-eswc-2017/

Cash prize!

Open Knowledge Extraction Challenge @ ESWC'17

When Simon & Garfunkel split in 1970, Simon quickly began his solo career with the release of the self-titled album "Paul Simon". This was followed by "There Goes Rhymin' Simon" and "Still Crazy After All These Years", both of which featured chart-topping hits such as "Loves Me Like A Rock" and "Kodachrome".

identified named entity	classified type	generated URI	indices
Simon & Garfunkel	MusicArtist	artist:5d02f264-e225-41ff-83f7-d9b1f0b1874a	5,22
Simon	MusicArtist	artist:fc0a5289-4b77-4246-9c8d-857c8b617f5d	38,43
Paul Simon	SignalGroup	release-group:a1cc3fbd-609b-323c-95e2-435dfceb51e9	117,127
There Goes Rhymin' Simon	SignalGroup	release-group:fb1e90a8-4461-382b-9081-183abb3c8997	152,176
Still Crazy After All These Years	SignalGroup	release-group:cd0c17f4-ff8d-3b1d-ac36-397ebbb069e9	183,216
Loves Me Like A Rock	MusicalWork	work:bc76594b-b113-4a57-b929-b9911531108e	270,290
Kodachrome	MusicalWork	work:c9ad17e6-440e-40b6-b4f9-58b74b006c20	297,307

Future

Chatbots

Deep Learning + Semantics

Multimodality

Deep Generative Models

- Text generation from audio
- Audio generation from text

Thanks!

Questions? Ideas? Suggestions?

@sergiooramas @luisanke