

Semantic Relation Extraction. Resources, Tools and Strategies

Marcos García González

Centro Singular de Investigación en Tecnoloxías da Información

Universidade de Santiago de Compostela

June 7, 2013

citius.usc.es

Outline

1 Context and Motivation

2 Hypotheses and Objectives

3 Methodology

- (I) Natural Language Processing Tools
- (II) Relation Extraction

4 Results

5 Current and Further Work



Outline

1 Context and Motivation

2 Hypotheses and Objectives

3 Methodology

- (I) Natural Language Processing Tools
- (II) Relation Extraction

4 Results

5 Current and Further Work



What is Relation Extraction? And why doing it?



- Subtask of Information Extraction: obtain knowledge from unstructured text
- Useful for: Question Answering, Information Retrieval, Semantic Text Enrichment...

What kind of information?

■ Relation triples

`predicate(Arg1, Arg2)`

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés. Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés. Foi un dos máis prestixiosos científicos galegos dese século.”

hasAlias(Xosé Rodríguez González, Matemático Rodríguez)
hasAlias(Xosé Rodríguez González, Matemático de Bermés)
hasBirthDate(Xosé Rodríguez González, 05/07/1770)
hasBirthPlace(Xosé Rodríguez González, Bermés)
hasProfession(Xosé Rodríguez González, científico)

What kind of information?

■ Relation triples

`predicate(Arg1, Arg2)`

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés. Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés. Foi un dos máis prestixiosos científicos galegos dese século.”

`hasAlias(Xosé Rodríguez González, Matemático Rodríguez)`

`hasAlias(Xosé Rodríguez González, Matemático de Bermés)`

`hasBirthDate(Xosé Rodríguez González, 05/07/1770)`

`hasBirthPlace(Xosé Rodríguez González, Bermés)`

`hasProfession(Xosé Rodríguez González, científico)`

What kind of information?

■ Relation triples

`predicate(Arg1, Arg2)`

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés. Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés. Foi un dos máis prestixiosos científicos galegos dese século.”

`hasAlias(Xosé Rodríguez González, Matemático Rodríguez)`

`hasAlias(Xosé Rodríguez González, Matemático de Bermés)`

`hasBirthDate(Xosé Rodríguez González, 05/07/1770)`

`hasBirthPlace(Xosé Rodríguez González, Bermés)`

`hasProfession(Xosé Rodríguez González, científico)`

What kind of information?

■ Relation triples

`predicate(Arg1, Arg2)`

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés. Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés. Foi un dos máis prestixiosos científicos galegos dese século.”

`hasAlias(Xosé Rodríguez González, Matemático Rodríguez)`

`hasAlias(Xosé Rodríguez González, Matemático de Bermés)`

`hasBirthDate(Xosé Rodríguez González, 05/07/1770)`

`hasBirthPlace(Xosé Rodríguez González, Bermés)`

`hasProfession(Xosé Rodríguez González, científico)`

Practical Example

■ Knowledge Graph

Web Imagens Mapas Compras Notícias Mais ▾ Ferramentas de pesquisa

Cerca de 2 110 000 resultados (0,19 segundos)

Josep Guardiola – Wikipédia, a encyclopédia livre
pt.wikipedia.org/wiki/Josep_Guardiola ▾ Traduzir esta página

Josep Guardiola | Sala, também conhecido como Pep (Santpedor, 18 de janeiro de 1971), é um ex-futebolista, e futuro treinador do Bayern de Munique para a ...
 Carreira - Como jogador - Títulos - Referências

Pep Guardiola - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Pep_Guardiola ▾ Traduzir esta página

José "Pep" Guardiola | Sala é a Spanish football manager who will become manager of German Bundesliga club Bayern Munich at the start of the 2013–14 ...

Pep Guardiola – Wikipedia, la encyclopédie libre
es.wikipedia.org/wiki/Pep_Guardiola ▾ Traduzir esta página

Josep Guardiola | Sala [ʒu'zɛp gwər'diɔla] (Sampedor, Barcelona, 18 de enero de 1971), conocido como Pep Guardiola o simplemente Guardiola, es un ...

Página Web de Josep Guardiola
<www.josepguardiola.com/> ▾ Traduzir esta página

guardiola, Guardiola, Josep, josep, Josep Guardiola, Pep Guardiola, pep guardiola fotos, Josep Guardiola fotos, pep guardiola fotos, Pep Guardiola fotos, Barça ...

Notícias sobre Josep guardiola

José Mourinho y Josep Guardiola se enfrentarán por la Supercopa de Europa | Peru.com
 Peru.com - Há 1 dia
 Ambos técnicos ya tienen fecha, hora y estadio pactado para volverse a ver en una cancha. Esta vez chocarán fuera de los buzos del ...

Pep Guardiola ignora cumprimento de Cristiano Ronaldo em festa ...
www.youtube.com/watch?v=IMDokk_WRUk
 09/01/2013 - Carregado por DILMILLIIVULGA :D
 O ex-técnico do Barcelona, Josep GUARDIOLA, é idolatrado por torcedores em todo mundo, mas reagiu ...

Mais vídeos de **josep guardiola** »



Josep Guardiola

José Guardiola | Sala, também conhecido como Pep, é um ex-futebolista, e futuro treinador do Bayern de Munique para a temporada 2013-2014. Wikipedia

Nascimento: 18 de janeiro de 1971 (42 anos), Santpedor

Altura: 1,80 m

Cônjuge: Cristina Serra

Filhos: Valentina Guardiola, Marius Guardiola, Maria Guardiola

Resquícios relacionados



Reportar erro/Mais informações

Outline

1 Context and Motivation

2 Hypotheses and Objectives

3 Methodology

- (I) Natural Language Processing Tools
- (II) Relation Extraction

4 Results

5 Current and Further Work



Hypothesis and Objectives

Hypothesis

Is it possible to automatically extract instances of semantic relations accurately?

Objective

To analyze and assess different Relation Extraction (RE) strategies for PT/GL/ES (namely encyclopaedic knowledge)

Objective II

To evaluate Natural Language Processing (NLP) tools and resources required for RE systems (and create them, if necessary)

Hypothesis and Objectives

Hypothesis

Is it possible to automatically extract instances of semantic relations accurately?

Objective

To analyze and assess different Relation Extraction (RE) strategies for PT/GL/ES (namely encyclopaedic knowledge)

Objective II

To evaluate Natural Language Processing (NLP) tools and resources required for RE systems (and create them, if necessary)

Outline

1 Context and Motivation

2 Hypotheses and Objectives

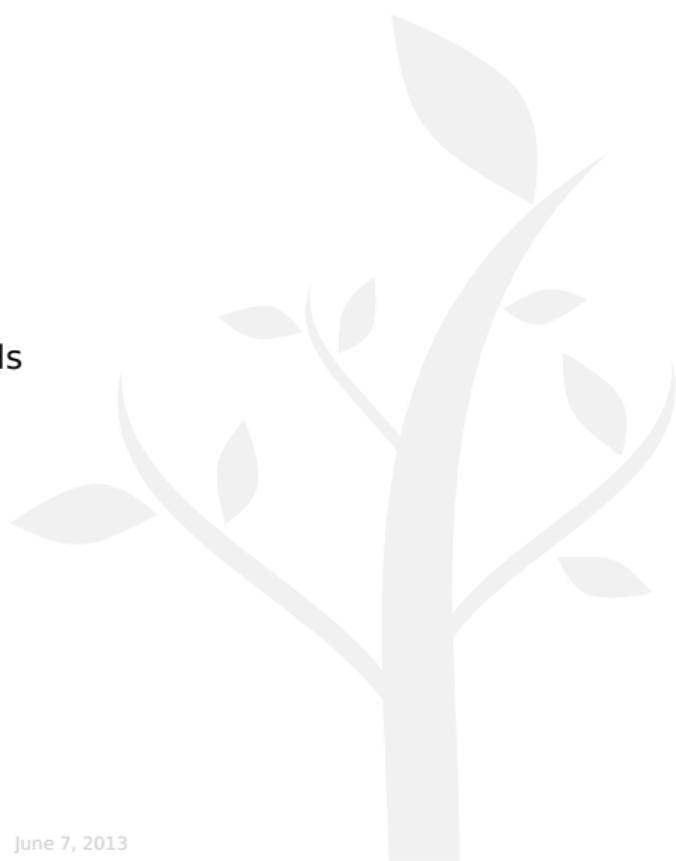
3 Methodology

(I) Natural Language Processing Tools

(II) Relation Extraction

4 Results

5 Current and Further Work



Part I

Natural Language Processing Tools

Text Pre-processing

- Previous linguistic processing in order to make the extraction easier
- Several NLP tasks
- Mostly performed with FreeLing (open-source, multilingual)

Sentence Splitter and Tokenization

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés. Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés. Foi un dos máis prestixiosos científicos galegos dese século.”

- Sentence Splitter (FST)
- Tokenization (dictionary and rule-based)

Sentence Splitter and Tokenization

“<s> Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés.</s> <s> Naceu o 5 de xullo de 1770 na parroquia lalinense de Bermés.</s> <s> Foi un dos máis prestixiosos científicos galegos dese século.</s> ”

- Sentence Splitter (FST)
- Tokenization (dictionary and rule-based)

Sentence Splitter and Tokenization

“<s>Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . </s><s> Naceu o 5 de xullo de 1770 **en a** parroquia lalinense de Bermés . </s><s> Foi un **de os** más prestixiosos científicos galegos **de ese** século . </s>”

- Sentence Splitter (FST)
- Tokenization (dictionary and rule-based)

Lemmatization and Part-of-Speech (PoS) Tagging

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . Naceu o 5 de xullo de 1770 en a parroquia lalinense de Bermés . Foi un de os más prestixiosos científicos galegos de ese século .”

Lemmatization (dict. and rules)

é → ser

coñecido → coñecer

Naceu → nacer

Foi → ser

PoS-tagging (HMM)

é → Verb

como → Conjunction

Matemático → Proper Noun

un → Determiner

Lemmatization and Part-of-Speech (PoS) Tagging

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . Naceu o 5 de xullo de 1770 en a parroquia lalinense de Bermés . Foi un de os más prestixiosos científicos galegos de ese século .”

Lemmatization (dict. and rules)

é → ser

coñecido → coñecer

Naceu → nacer

Foi → ser

PoS-tagging (HMM)

é → Verb

como → Conjunction

Matemático → Proper Noun

un → Determiner

Lemmatization and Part-of-Speech (PoS) Tagging

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . Naceu o 5 de xullo de 1770 en a parroquia lalinense de Bermés . Foi un de os más prestixiosos científicos galegos de ese século .”

Lemmatization (dict. and rules)

é → ser

coñecido → coñecer

Naceu → nacer

Foi → ser

PoS-tagging (HMM)

é → Verb

como → Conjunction

Matemático → Proper Noun

un → Determiner

Named Entity Recognition

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . Naceu o 5 de xullo de 1770 en a parroquia lalinense de Bermés . Foi un de os máis prestixiosos científicos galegos de ese século .”

- Named Entity and Dates Recognition (FST and Viterbi)
- Semantic Classification (AdaBoost and resource-based)

Named Entity Recognition

“Xosé Rodríguez González é tamén coñecido como Matemático Rodríguez ou o Matemático de Bermés . Naceu o 5 de xullo de 1770 en a parroquia lalinense de Bermés . Foi un de os máis prestixiosos científicos galegos de ese século .”

- Named Entity and Dates Recognition (FST and Viterbi)
- Semantic Classification (AdaBoost and resource-based)

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession **gobernador de Texas**

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession **gobernador de Texas**
Bush hasFather **George H.W. Bush**

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

Bush hasFather George H.W. Bush

Bush hasBrother Jeb Bush

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

Bush hasFather George H.W. Bush

Bush hasBrother Jeb Bush

Person hasBirthPlace New Haven

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

Bush hasFather George H.W. Bush

Bush hasBrother Jeb Bush

Person hasBirthPlace New Haven

George H. W. Bush hasWife Barbara

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

Bush hasFather George H.W. Bush

Bush hasBrother Jeb Bush

Person hasBirthPlace New Haven

George H. W. Bush hasWife Barbara

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

Bush hasFather George H.W. Bush

Bush hasBrother Jeb Bush

Person hasBirthPlace New Haven

George H. W. Bush hasWife Barbara

Co-reference Resolution

“George Walker Bush fue el cuadragésimo tercer presidente de los Estados Unidos de América [...]. George W. Bush fue elegido gobernador de Texas en 1994 [...]. Bush es hijo del expresidente George H. W. Bush , hermano del ex gobernador del estado de Florida Jeb Bush y nieto del ex senador Prescott Bush . Ø Nació en New Haven, Connecticut, siendo el primer hijo del matrimonio formado George H. W. Bush y su esposa Barbara .”

- Mentions (and zero-pronouns) referring to the same *real entity* (rule-based and distance-matching)

George W. Bush hasProfession gobernador de Texas

George W. Bush hasFather George H.W. Bush

George W. Bush hasBrother Jeb Bush

George W. Bush hasBirthPlace New Haven

George H. W. Bush hasWife Barbara

Part II

Relation Extraction

Strategies

- Rule-based
- Supervised models
- Distant-supervised models
- Open Information Extraction



Rule-based models

- Closed set of relations
- Hand-crafted rules
 - ▷ High precision
 - ▷ Low recall

hasBirthPlace: Person:Subject_nacer; Location:Circ_nacer_en;

hasDeathDate: Person:Subject_morrer; Date:Circ_morller_en;

Supervised models

- Closed set of relations
- Need annotated corpora:

Person é o autor de Book (isAuthorOf)

Person escreveu o livro Book (isAuthorOf)

Book foi terminado no ano Date por Person (isAuthorOf)

Person assinou exemplares do seu livro Book (isAuthorOf)

- High precision and recall (depending on the relation and on the training corpus)

Supervised models

- Closed set of relations
- Need annotated corpora:

Person é o autor de Book (isAuthorOf)

Person escreveu o livro Book (isAuthorOf)

Book foi terminado no ano Date por Person (isAuthorOf)

Person assinou exemplares do seu livro Book (isAuthorOf)

- High precision and recall (depending on the relation and on the training corpus)

Distant-supervised models

- ***Not so closed set of relations***

- Seed pairs (from databases: e.g., Freebase, DBpedia):

Pep Guardiola — Santpedor

Linus Torvalds — Helsinki

Ernest Hemingway — Oak Park

- Automatic building of annotated corpora:

Person nasceu na vila de Location (hasBirthPlace)

Person, nascido no ano Date em Location (hasBirthPlace)

Location é a cidade de nascimento do escritor Person (hasBirthPlace)

- May produce noisy sentences:

Person visitou uns amigos em Location *(hasBirthPlace)

Distant-supervised models

- *Not so closed set of relations*
- Seed pairs (from databases: e.g., Freebase, DBpedia):

Pep Guardiola — Santpedor

Linus Torvalds — Helsinki

Ernest Hemingway — Oak Park

- Automatic building of annotated corpora:

Person nasceu na vila de Location (hasBirthPlace)

Person, nascido no ano Date em Location (hasBirthPlace)

Location é a cidade de nascimento do escritor Person (hasBirthPlace)

- May produce noisy sentences:

Person visitou uns amigos em Location *(hasBirthPlace)

Distant-supervised models

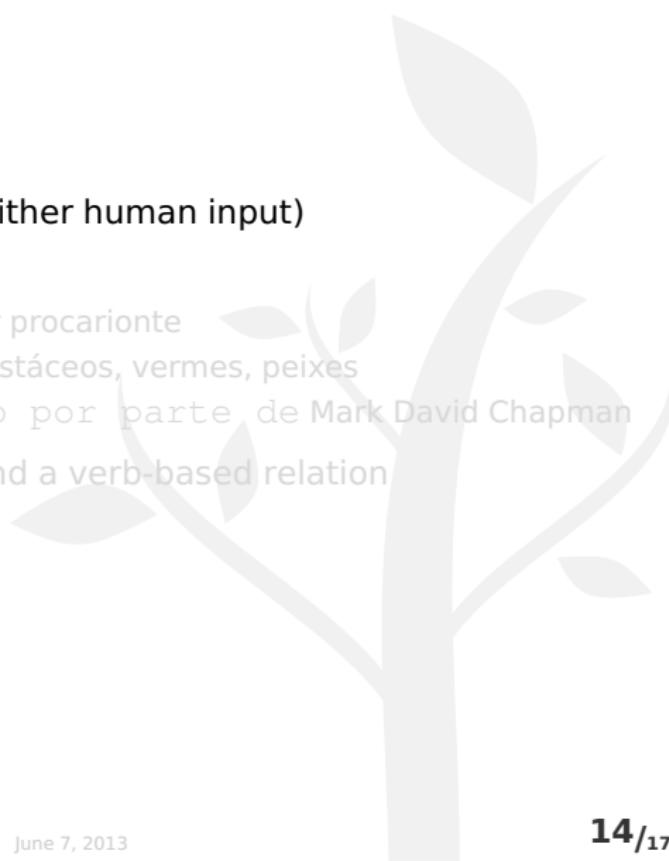
- *Not so closed set of relations*
- Seed pairs (from databases: e.g., Freebase, DBpedia):
 - Pep Guardiola — Santpedor
 - Linus Torvalds — Helsinki
 - Ernest Hemingway — Oak Park
- Automatic building of annotated corpora:
 - Person nasceu na vila de Location (hasBirthPlace)
 - Person, nascido no ano Date em Location (hasBirthPlace)
 - Location é a cidade de nascimento do escritor Person (hasBirthPlace)
- May produce noisy sentences:
 - Person visitou uns amigos em Location *(hasBirthPlace)

Distant-supervised models

- *Not so closed* set of relations
- Seed pairs (from databases: e.g., Freebase, DBpedia):
 - Pep Guardiola — Santpedor
 - Linus Torvalds — Helsinki
 - Ernest Hemingway — Oak Park
- Automatic building of annotated corpora:
 - Person nasceu na vila de Location (hasBirthPlace)
 - Person, nascido no ano Date em Location (hasBirthPlace)
 - Location é a cidade de nascimento do escritor Person (hasBirthPlace)
- May produce noisy sentences:
 - Person visitou uns amigos em Location *(hasBirthPlace)

Open Information Extraction

- Unlimited number of relations
- No need of annotated corpora (neither human input)
- Extracts billions of facts:
 - ▷ bacteria é organismo unicelular procarionte
 - ▷ robaliza aliméntase de crustáceos, vermes, peixes
 - ▷ John Lennon foi assassinado por parte de Mark David Chapman
- Triples between two arguments and a verb-based relation
(overspecification)



Open Information Extraction

- Unlimited number of relations
- No need of annotated corpora (neither human input)
- Extracts billions of facts:
 - ▷ **bacteria é organismo unicelular procarionte**
 - ▷ **robaliza aliméntase de crustáceos, vermes, peixes**
 - ▷ **John Lennon foi assassinado por parte de Mark David Chapman**
- Triples between two arguments and a verb-based relation
(overspecification)

Open Information Extraction

- Unlimited number of relations
- No need of annotated corpora (neither human input)
- Extracts billions of facts:
 - ▷ **bacteria** é organismo unicelular procarionte
 - ▷ **robaliza** alimentase de crustáceos, vermes, peixes
 - ▷ **John Lennon** foi assassinado por parte de **Mark David Chapman**
- Triples between two arguments and a verb-based relation
(overspecification)

Outline

1 Context and Motivation

2 Hypotheses and Objectives

3 Methodology

- (I) Natural Language Processing Tools
- (II) Relation Extraction

4 Results

5 Current and Further Work



Current Results

- Not directly comparable
- Vary greatly depending on the relation/corpus/language
 - ▷ Rule-based: > 95% precision and \approx 60% recall
 - ▷ Supervised: > 85% precision and \approx 80% recall
 - ▷ Distant-supervised: > 70% precision and \approx 60% recall
 - ▷ Open Information Extraction: > 65% precision and \approx 50% recall

Outline

1 Context and Motivation

2 Hypotheses and Objectives

3 Methodology

- (I) Natural Language Processing Tools
- (II) Relation Extraction

4 Results

5 Current and Further Work



What am I doing now?

- Improve co-reference resolution modules
- Evaluate how co-reference resolution works with Open Information Extraction
- Perform large-scale extractions in PT/GL/ES

Publications

Journal Papers:

- Garcia, Marcos and Pablo Gamallo, submitted. Exploring the Effectiveness of Linguistic Knowledge for Biographical Relation Extraction. *Natural Language Engineering*.
- Garcia, Marcos, Iria Gayo and Isaac González López, 2012. Identificación e Clasificación de Entidades Mencionadas em Galego. *Estudos de Lingüística Galega*, 4.
- Garcia, Marcos and Pablo Gamallo, 2011. Resolución de Correferencia de Nombres de Persona para Extracción de Información Biográfica. *Procesamiento del Lenguaje Natural*, 47.
- Garcia, Marcos and Pablo Gamallo, 2010. Análise Morfossintáctica para Português Europeu e Galego: Problemas, Soluções e Avaliação. *Linguamática*, 2(2).

Conference Papers:

- Gamallo, Pablo, Marcos Garcia and Santiago Fernández-Lanza, 2012. Dependency-Based Open Information Extraction. In Proceedings of the ROBUS-UNSUP 2012 (EACL 2012).
- Garcia, Marcos and Pablo Gamallo, 2011. A Weakly-Supervised Rule-Based Approach for Relation Extraction. In Proceedings of CAEPIA 2011. KEEOS Workshop.
- Gamallo, Pablo and Marcos Garcia, 2011. A Resource-Based Method for Named Entity Extraction and Classification. In Proceedings of EPIA 2011, Progress in Artificial Intelligence (LNCS/LNAI).
- Garcia, Marcos and Pablo Gamallo, 2011. Dependency-Based Text Compression for Semantic Relation Extraction. In Proceedings of IEKA 2011 (RANLP 2011).
- Garcia, Marcos and Pablo Gamallo, 2011. Evaluating Various Linguistic Features on Semantic Relation Extraction. In Proceedings of RANLP 2011.
- Garcia, Marcos and Pablo Gamallo, 2011. An Exploration of the Linguistic Knowledge for Semantic Relation Extraction in Spanish. In Proceedings of the Joint Workshop FAM-LbR/KRAQ'11 (IJCAI 2011).
- Garcia, Marcos and Pablo Gamallo, 2010. Using Morphosyntactic Post-processing to Improve POS-tagging Accuracy. In Proceedings of PROPOR 2010.

Thanks!!!