



Readers: Evaluation And DEvelopment of Reading Systems

Anselmo Peñas
CHIST-ERA meeting, 5 March 2014, Istanbul

Overview

1. Starting date
2. Scientific background
3. Challenges and expected results
4. Partners and their value to the consortium
5. Workplan
6. Management
7. Outputs
8. Valorisation

Starting date

- **January 1st, 2013**
 - Kick-off meeting, Madrid, January 22-23, 2013



- Different starting dates of partners ...
 - U. Edinburgh, UK
 - EPSRC EP/K017845/1
 - October 1st, 2012
 - Synapse Développement, France
 - ANR-12-CHRI-0004-03
 - November 1st, 2012
 - UNED and UPV/EHU, Spain
 - MINECO PCIN-2013-002-C02
 - July 1st, 2013

Scientific Background

- Learning by Reading (DARPA 2005)
 - Pilot studies
 - Semantic representation of documents
 - Aligned to Knowledge Bases
 - Bottleneck: Knowledge Integration
- Machine Reading Program (DARPA 2009)
 - “Develop technology enabling formal processing over naturally occurring text”
 - Decouple Reading and Reasoning machines
 - Limitation of classic supervised Information Extraction
- Leverage redundancy: from local to global
 - Knowledge Bases cannot grow on their own
 - Develop semi-supervised and unsupervised approaches
 - Macro-reading
 - Open Information Extraction



Scientific Background

Attached to a KB

- Micro-reading take advantage of the KB
 - Known instances and their categories
 - Subset /superset relations
 - Typed relations
 - Constraints
- Macro-reading populates the KB
 - Semi-supervised relation extraction + Entity Linking
 - Leverage redundancy
 - Beyond the sum of micro-readings
 - Self-growing Relational Knowledge Bases
- At the same time is continuously increasing its ability to map
 - More seeds -> more patterns
 - More patterns -> more seeds

Scientific Background

- But imagine you are able to get all patterns...
- Would it be enough?
- A single message is never self-contained
- There is always implicit information

$$\frac{\text{New_Information}}{\text{Background_Knowledge_required}} \sim 0$$



Scientific Background

- Is this implicit information expressable by our KBs?
 - Default answer is: No
- Then we can't use our KB to make explicit the implicit information

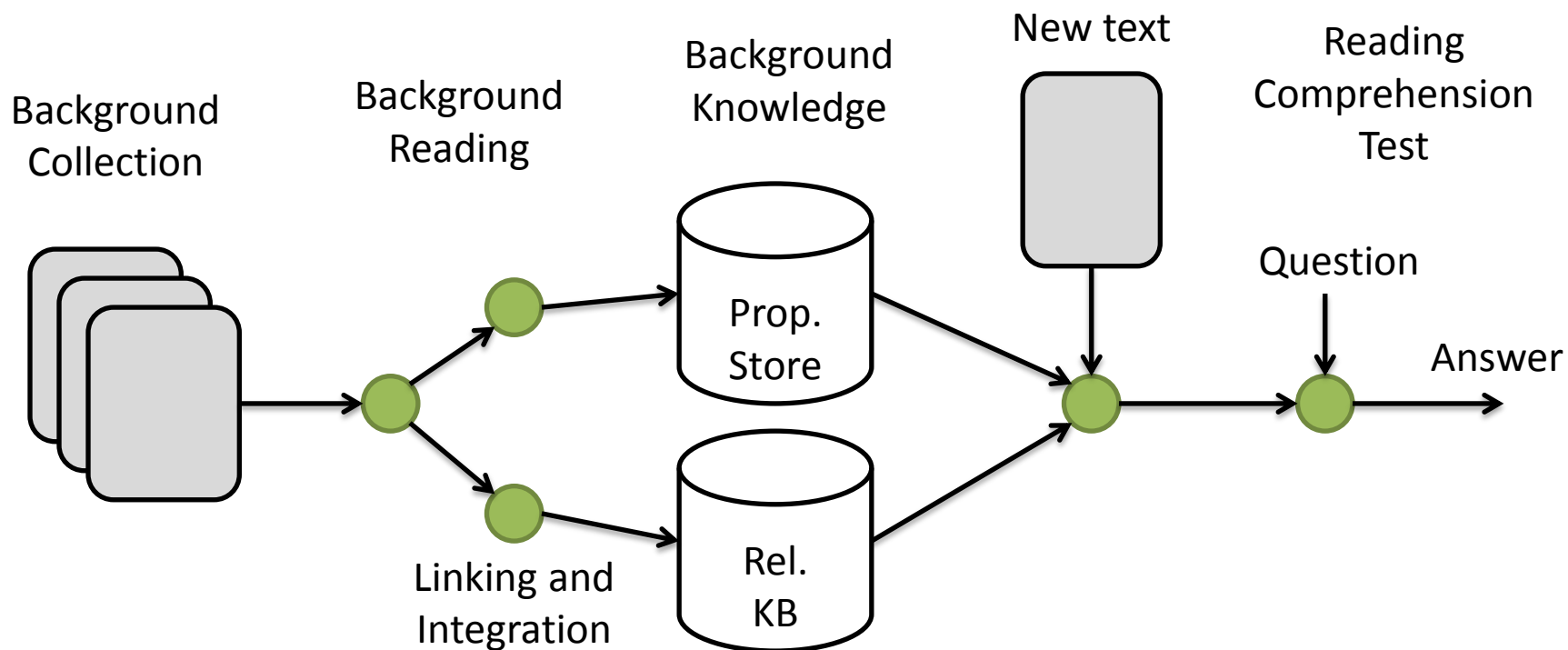
And what is missing...

... can't be used by a pattern to extract information

- We need to explore additional alternatives
 - Not attached to predefined relational KBs
 - Unsupervised

Scientific Background

● Reading Technology (READERS)





Challenges

- Background Reading
 - Induction of semantic roles
 - Build propositional background knowledge bases
- Knowledge Linking and Integration
 - Acquired background knowledge
 - Existing Knowledge Bases
- Read a single document
 - Bringing the background knowledge
 - Required to answer questions about the doc.
- Develop evaluation methodologies for Machine Reading

Expected Results

- Advance the state of the art in Machine Reading and Natural Language Understanding
- New unsupervised and distant supervised methods for:
 - background knowledge acquisition
 - semantic role induction
 - automatic taxonomy creation
 - discovery of class-instance relations.
- New propositional knowledge bases linking them to existing knowledge repositories

Expected Results

- Algorithms for efficient storage and use of the knowledge resulting from processing large amounts of unstructured text
- A real industrial Machine Reading system
- Protocols for measuring the progress in the development of Machine Reading technology
- Organization of an international campaign for Machine Reading evaluation in several languages, with 100% reusable benchmarks by research groups and companies.
- Demonstrators of the technology for specific customer real use cases



Consortium

Country	Institution/Department	Principal Investigator	Co-investigators
Spain	UNED (coordination)	Anselmo Peñas	Henry Anaya
United Kingdom	U. Edinburgh	Maria Mirella Lapata	Kristian Woodsend
France	Synapse Développement	Patrick Séguéla	Dominique Laurent Baptiste Chardon
Spain	UPV/EHU	Eneko Agirre	German Rigau

Consortium

UNED NLP & IR Group, Spain

- 10 senior PhD
- 12 researchers/students

— Research lines:

- Multilingual Multimodal Information Access
- Text Mining
- Online Reputation Management
- Information Extraction and Textual Inference
- **Information Access Evaluation**

Consortium

Institute for Language, Cognition, and
Computation (ILCC), U. Edinburgh, UK

- 27 academic faculty
- 78 researchers/students

— Research lines:

- Linguistic resources and tools
- Information Extraction
- Statistical Machine Translation
- Large-scale robust NLP
- **Probabilistic learning for NL Understanding**
- **Extracting semantic information from large volumes of text**
- **Unsupervised methods**

Consortium

Synapse Développement, France

- Company specialized in linguistic software development
- Founded in 1994
- B2C: more than 30,000 clients
- B2B

— Research lines:

- Linguistic analyzers
- Opinion detection
- **Question Answering**

Consortium

IXA NLP group, Spain

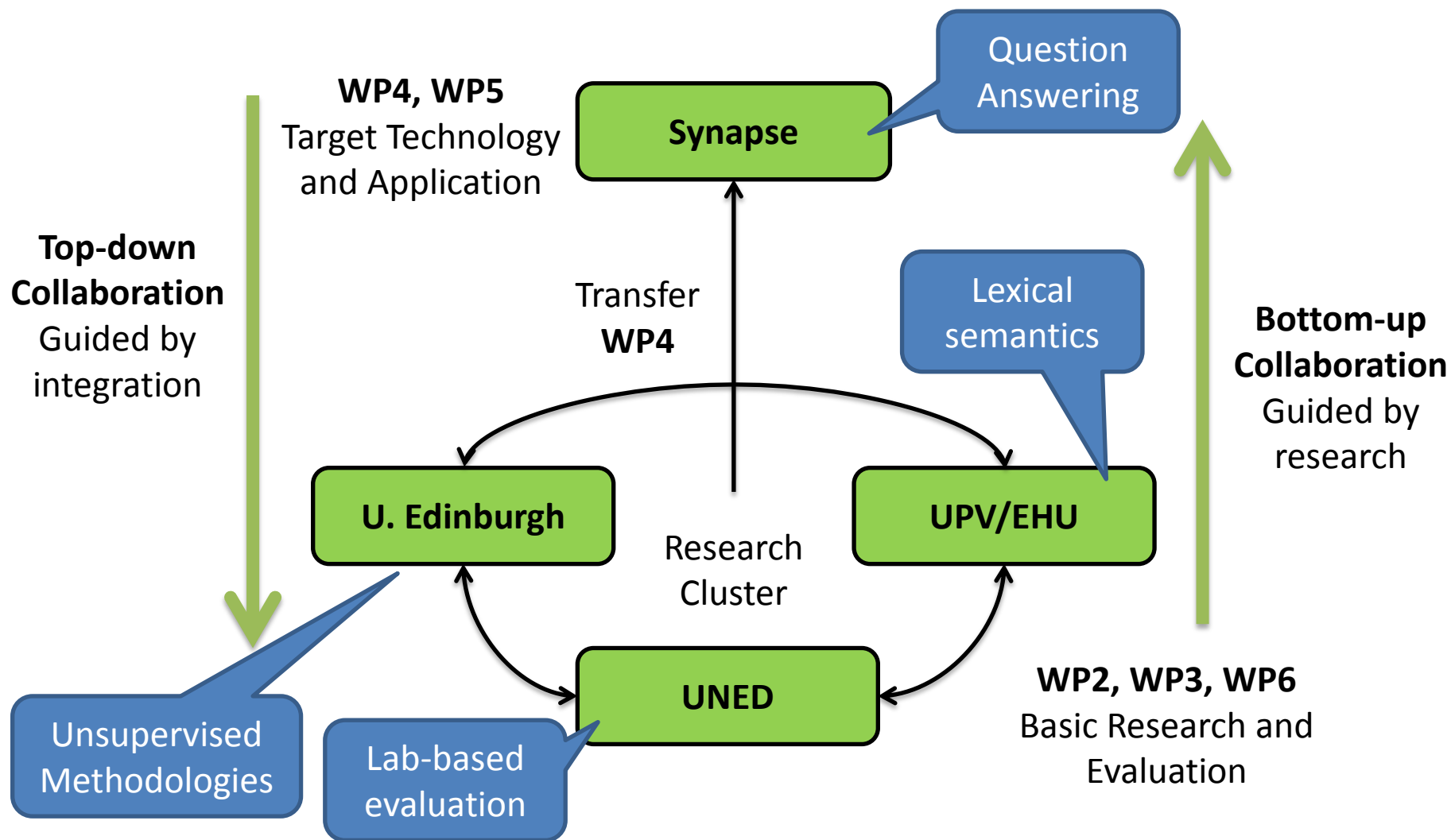
Interdisciplinary team

- 33 CS and 15 linguists
- 28 PhD
- 26 students/research assistants

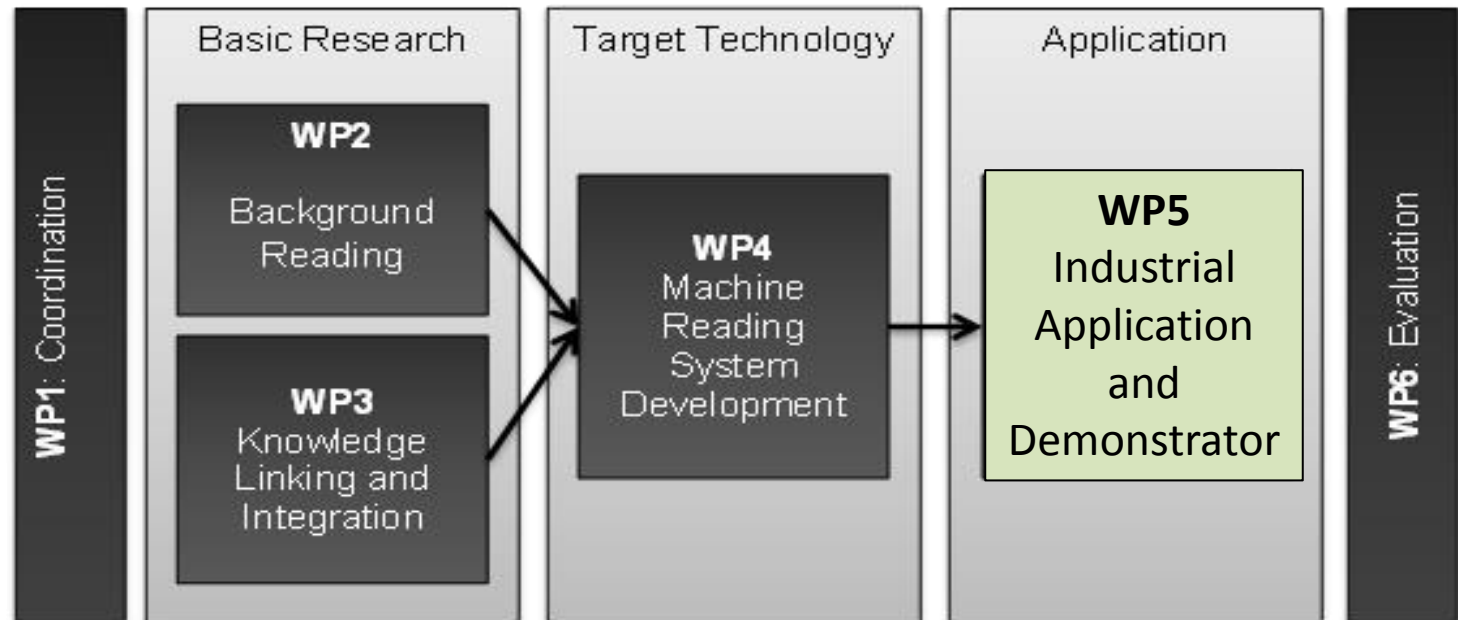
— Research lines:

- Building NLP resources and tools for Basque
 - Also English and Spanish
- Applications: IR, IE, MT, education
- Commercial: spell-checker, IR, dictionaries, etc.
- ... **lexico-semantic resources**
- ... **semantic processing**

Consortium



Workplan and deviation



person-month	WP1	WP2	WP3	WP4	WP5	WP6	total
1 (UNED)	8	12	1	2	1	12	36
2 (U. Edinburgh)	2	20	6	4	2	2	36
3 (Synapse Développement)	2	0.5	0.5	66	20	3	92
4 (UPV/EHU)	2	1	18	1	1	1	24
total	14	33.5	25.5	73	24	18	188

WP1 Coordination

	Task leader
WP1 Coordination	UNED
1.1 Work plan monitoring	UNED
1.2 Reporting and Dissemination	UNED
1.3 Collaboration with third parties and Community building	UNED

	Deliverable	Month of delivery	Title of deliverable
✓	D1.1	M12	1 st annual report
	D1.2	M24	2 nd annual report
	D1.3	M36	Final report

WP2 Background Reading

	Task leader
WP2 Background Reading	ED
2.1 Processing of document collections	UNED
2.2 Generation of Proposition Stores	UNED
2.3 Robust Semantic Role Labeling (SRL)	ED
2.4 SRL for multiple languages	ED
2.5 SRL without parsing	ED
2.7 Learning to read	UPV/EHU
2.8 Background Knowledge Base creation	ED

Deliverable	Month of delivery	Title of deliverable
✓ D2.1	M12	Initial Proposition Store
D2.2	M24	Semantic Role Labelling component
D2.3	M36	Open IE and knowledge base construction

WP3 Knowledge Linking and Integration

	Task leader
WP3 Knowledge Linking and Integration	UPV/EHU
3.1 Preparation of Knowledge Bases	UPV/EHU
3.2 KB interface for disambiguation	UPV/EHU
3.3 KB interface for relation Extraction	UPV/EHU
3.4 KB interface for similarity, relatedness and expansion	UPV/EHU
3.5 Knowledge Base population	UPV/EHU
3.6 Taxonomy Induction	ED

Deliverable	Month of delivery	Title of deliverable
✓ D3.1	M6	Design of knowledge linking and integration
D3.2	M30	Results of knowledge linking and integration

WP4 Machine Reading System Development

	Task leader
WP4 Machine Reading System Development	SD
4.1 Development of Multilingual platform	SD
4.2 Development of French Reading Machine	SD
4.3 Development of English Reading Machine	SD
4.4 Software release	ED

Deliverable	Month of delivery	Title of deliverable
✓ D4.1	M12	Storing format in Knowledge Base
D4.2	M24	Multilingual platform with beta-version
D4.3	M24	SRL Software release
D4.4	M36	Multilingual platform with final version
D4.5	M36	Taxonomy induction software release



New WP5 (Industrial Application)

- Application initially planned:
 - Automatic Generation of Reading Tests to assess children attitude
- Synapse evaluated the commercial potential of this application
 - Results showed difficulties for a company to get outcomes out of it
- Re-shape WP5 as described in D5.0
 - Starting with the construction of real use cases
 - After interviews with companies with potential interest in the technology
- Better strategy
 - Customer oriented, higher value for Synapse
 - EADS – Airbus use case (Eurocopter, England)
 - Help “Eurocopter” teams in the management of workflows (textual messaging based) for maintenance support

WP5 Industrial Application and Demonstrator

	Task leader
WP5 Industrial Application and Demonstrator	SD
5.1 Build use cases of interest for the industry	SD
5.2 Design the demonstrator of the Readers technology	SD
5.3 Demonstrator Release	SD

Deliverable	Month of delivery	Title of deliverable
✓ D5.0	M6	New definition of WP5
✓ D5.1	M12	Report on building comprehensive business case
D5.2	M24	Report on the design of the demonstrator and the use case
D5.3	M32	Demonstrator Release

WP6 Evaluation of Machine Reading

	Task leader
WP6 Evaluation of Machine Reading	UNED
6.1 Definition of the campaign exercises and goals	UNED
6.2 Evaluation methodology and resources	UNED
6.3 Activation of the campaign exercises	UNED
6.4 Assessment and comparison of systems results	UNED
6.5 Workshop	UNED

Deliverable	Month of delivery	Title of deliverable
✓ D6.1	M12	1 st Report on the lab-based evaluation
D6.2	M24	2 nd Report on the lab-based evaluation
D6.3	M36	3 rd Report on the lab-based evaluation

Milestones

No of Milestone	Delivery month	WP involved	Title
✓ M1	M12	WP2	Initial Proposition Store
M2	M24	WP2, WP4	Semantic Role Labelling Software release
M3	M36	WP3, WP4, WP6	Machine Reading system
M4	M36	WP4, WP5	Industrial demonstrator release

Management – meetings

- Kick-off meeting, Madrid, January 22-23, 2013
- Annual meeting, San Sebastian, January 21-22, 2014
- Monthly phone conference of the Technical Coordination Committee (TCC)
 - Integrated by Work Package Leaders (WPLs)

N°	Date	Location	Attending partners	Purpose
1	22-23/1/2013	Madrid	UNED, ED, Synapse, UPV/EHU	Kick-off meeting
2	22/2/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
3	15/3/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
4	26/4/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
5	31/5/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
6	27/6/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
7	13/9/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
8	13/12/2013	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Progress report by WP Leaders
9	10/1/2014	Phone Conf.	UNED, ED, Synapse, UPV/EHU	Preparation of annual meeting
10	21-22/1/2014	San Sebastian	UNED, ED (via Skype), Synapse, UPV/EHU	Annual meeting

Management – collaboration

- Exchange
 - Bernardo Cabaleiro (UNED)
 - Internship in Edinburgh: 10/9/2012 – 10/12/2012
 - Planning a second stay in 2014
 - Oier Lopez de Lacalle (UPV/EHU)
 - Postdoc at Edinburgh until 2014
- Creation of a new evaluation space, namely CLEF QA Track 2014 (with QALD and BioASQ)
- Coordination with the Japanese National Institute of Informatics (Prof. Noriko Kando) to co-organize the “Entrance Exams” challenge

Management – financial

Nº	Partner	Person.months	Total costs (Euros)	Percentage of requested budget
1	UNED	6	0	0 %
2	U. Edinburgh	12	127,000	33 %
3	Synapse	22	130,882	25.36 %
4	UPV/EHU	6	0	0 %

Note:

- UNED and UPV/EHU received funds at the end of the period (December 2013)
- Personnel budget only for new contracts

Management – web site

Website: nlp.uned.es/readers-project/



Readers: Evaluation And DEvelopment of Reading Systems

[Summary](#) - [Objectives](#) - [Partners](#)

Summary

The READERS project proposes new unsupervised computational models to automatically extract background knowledge after reading large amounts of unstructured text. This knowledge will be in the form of classes, categorized entities and predicates whose arguments are typified by probability distributions over classes. Classes themselves will be automatically organized into taxonomies related to the predicates in which they participate.

In this way, new methods and models based on extensional definitions of concepts will be developed and deployed for the automatic creation of knowledge bases. Importantly, these will be closely related to textual representations and instrumental in enabling textual inferences. The extracted knowledge will be also linked to external human-made resources such as Freebase, DBPedia and WordNet, and the knowledge bases will be interfaced with several engines for performing disambiguation, relation extraction, term expansion, and measuring relatedness.

A key part of the project will be the development of a reading matching that will use all these resources and tools. The purpose of our reading machine is to answer queries about a given text. Texts are never self-contained and their interpretation always requires recovering large amounts of background knowledge. Thus, the Machine Reading technology under development must incorporate not only language processing but also the recovery and use of large amounts of background knowledge.

This Machine Reading technology will be evaluated through Multiple-Choice Reading Comprehension tests (MRC) developed by humans over unseen documents. MRC tests enable objective and reproducible evaluation experiments, and will be 100% reusable as benchmarks available for the international community. Interestingly, the industrial partner in charge of the Machine Reading system development will apply the reverse technology to automatically generate MRC tests for the automatic assessment of children's reading abilities. This reading machine will work with at least two languages, English and French.

Outputs

1. Initial Proposition Store
 - Used to automatically unfold transfers of meaning in eventive propositions (CICLING 2014)
2. A novel deterministic algorithm for implicit Semantic Role Labeling (ACL 2013)
3. Predicate Matrix, a new lexical resource resulting from the integration of FrameNet, VerbNet, PropBank and WordNet (GWC 2014)
4. Development of systems for Entity Linking, Slot Filling and Temporal Slot Filling (TAC KBP 2013)

Outputs

5. Development of topic model based methods for filtering of tweets (CLEF RepLab 2013)
6. General coordination of the Question Answering for Machine Reading (CLEF QA4MRE 2013)
7. Organization of the Entrance Exams exercise on Reading Comprehension (CLEF QA4MRE 2013)
 - In collaboration with NTCIR and Todai Robot project (Japanese National Institute of Informatics)
8. Organization of *SEM 2013 shared task on Semantic Textual Similarity (STS)

Publications – methods

Multiple partners international conferences

1. Anselmo Peñas, Bernardo Cabaleiro and Mirella Lapata. **Unsupervised Interpretation of Eventive Propositions**. In Gelbukh (Ed.): CICLing 2014, Lecture Notes in Computer Science. Springer-Verlag 2014 (In Press)
2. Intxaurreondo A., Surdeanu M., Lopez de Lacalle O. and Agirre E. **Removing Noisy Mentions for Distant Supervision**. Procesamiento del Lenguaje Natural, 51, 41-48. 2013. (READERS mention is missing)
<http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/4870/2885>

Single partner international conferences

1. Laparra E. and Rigau G. **ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling**. The 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013). Sofia, Bulgaria, 2013. pp. 1180-1189.
<http://adimen.si.ehu.es/~rigau/publications/acl13-lr.pdf>
1. Henry Anaya-Sanchez, Anselmo Peñas, and Bernardo Cabaleiro. **UNED-READERS: Filtering Relevant Tweets using Probabilistic Signature Models**. Pamela Forner, Roberto Navigli, Dan Tufis (Eds.) CLEF 2013 Evaluation Labs and Workshop. Online Working Notes, ISBN 978-88-904810-5-5. ISSN 2038-4963. <http://www.clef-initiative.eu/documents/71612/35326d07-6b2d-4f41-bb7c-a93858f27470>

Publications – systems & resources

1. Barrena A., Agirre E. and Soroa A. **UBC Entity Linking at TAC-KBP 2013: random forests for high accuracy.** Text Analysis Conference (TAC 2013), Gaithersburgh, US, 2013.
<http://ixa.si.ehu.es/ixa/Argitalpenak/Artikuluak/1391179667/publikoak/tac2013>
1. Guillermo Garrido, Anselmo Peñas and Bernardo Cabaleiro. **UNED Slot Filling and Temporal Slot Filling systems at TAC KBP 2013. System description.** Text Analysis Conference (TAC 2013), Gaithersburgh, US. (In press)
2. López de Lacalle M., Laparra E. and Rigau G. **First steps towards a Predicate Matrix.** Proceedings of the 7th Global WordNet Conference (GWC 2014). Tartu, Estonia, 2014. <http://adimen.si.ehu.es/~rigau/publications/gwc14-llr.pdf>

Publications – evaluation

1. Anselmo Peñas, Eduard Hovy, Pamela Forner, Alvaro Rodrigo, Richard Sutcliffe, Roser Morante. QA4MRE 2011-2013: **Overview of Question Answering for Machine Reading Evaluation**. Information Access Evaluation. Multilinguality, Multimodality, and Visualization. Lecture Notes in Computer Science Volume 8138, Springer-Verlag 2013, pp. 303-320
http://dx.doi.org/10.1007/978-3-642-40802-1_29
2. Richard Sutcliffe, Anselmo Peñas, Eduard Hovy, Pamela Forner, Alvaro Rodrigo, Corina Forascu, Yassine Benajiba and Petya Osenova. **Overview of QA4MRE Main Task at CLEF 2013**. Pamela Forner, Roberto Navigli, Dan Tufis (Eds.) CLEF 2013 Evaluation Labs and Workshop. Online Working Notes, ISBN 978-88-904810-5-5. ISSN 2038-4963. <http://www.clef-initiative.eu/documents/71612/f2d296e0-ba40-4937-93bb-dcd0a22e9f4c>
3. Anselmo Peñas, Yusuke Miyao, Eduard Hovy, Pamela Forner and Noriko Kando. **Overview of QA4MRE 2013 Entrance Exams Task**. Pamela Forner, Roberto Navigli, Dan Tufis (Eds.) CLEF 2013 Evaluation Labs and Workshop. Online Working Notes, ISBN 978-88-904810-5-5. ISSN 2038-4963. <http://www.clef-initiative.eu/documents/71612/d3884d86-1876-44c3-8c93-a78e1168bd25>
1. Agirre E., Cer D., Diab M., Gonzalez-Agirre A., Guo W. ***SEM 2013 shared task: Semantic Textual Similarity**. The Second Joint Conference on Lexical and Computational Semantics (*SEM 2013), 2013. <http://ixa.si.ehu.es/ixa/Argitalpenak/Artikuluak/1371562669/publikoak/starsem2013-sharedtask>

Valorization

- Official collaboration between Synapse and **Airbus** in the framework of READERS project
 - One person of Airbus working on READERS
 - to propose use cases and assess viability
 - This person presented READERS project to **Airbus** to initiate official collaboration
- Synapse is developing the core technology for future demonstrators to its customers
- UPV/EHU delivers a disambiguation software
 - Continously updated including result of READERS project activity (<http://ixa2.si.ehu.es/ukb>)



Readers: Evaluation And DEvelopment of Reading Systems

Anselmo Peñas
CHIST-ERA meeting, 5 March 2014, Istanbul