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
• 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

Background Collection

Background Reading

Background Knowledge

Prop. Store

Rel. KB

Linking and Integration

New text

Question

Reading Comprehension Test

Answer

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

<table>
<thead>
<tr>
<th>Country</th>
<th>Institution/Department</th>
<th>Principal Investigator</th>
<th>Co-investigators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>UNED (coordination)</td>
<td>Anselmo Peñas</td>
<td>Henry Anaya</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>U. Edinburgh</td>
<td>Maria Mirella Lapata</td>
<td>Kristian Woodsend</td>
</tr>
<tr>
<td>France</td>
<td>Synapse Développement</td>
<td>Patrick Séguéla</td>
<td>Dominique Laurent Baptiste Chardon</td>
</tr>
<tr>
<td>Spain</td>
<td>UPV/EHU</td>
<td>Eneko Agirre</td>
<td>German Rigau</td>
</tr>
</tbody>
</table>
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
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
Workplan and deviation

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

<table>
<thead>
<tr>
<th>WP1 Coordination</th>
<th>Task leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Work plan monitoring</td>
<td>UNED</td>
</tr>
<tr>
<td>1.2 Reporting and Dissemination</td>
<td>UNED</td>
</tr>
<tr>
<td>1.3 Collaboration with third parties and Community building</td>
<td>UNED</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1.1</td>
<td>M12</td>
<td>1&lt;sup&gt;st&lt;/sup&gt; annual report</td>
</tr>
<tr>
<td>D1.2</td>
<td>M24</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt; annual report</td>
</tr>
<tr>
<td>D1.3</td>
<td>M36</td>
<td>Final report</td>
</tr>
</tbody>
</table>
### WP2 Background Reading

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

### Deliverables

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2.1</td>
<td>M12</td>
<td>Initial Proposition Store</td>
</tr>
<tr>
<td>D2.2</td>
<td>M24</td>
<td>Semantic Role Labelling component</td>
</tr>
<tr>
<td>D2.3</td>
<td>M36</td>
<td>Open IE and knowledge base construction</td>
</tr>
</tbody>
</table>
## WP3 Knowledge Linking and Integration

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3.1</td>
<td>M6</td>
<td>Design of knowledge linking and integration</td>
</tr>
<tr>
<td>D3.2</td>
<td>M30</td>
<td>Results of knowledge linking and integration</td>
</tr>
</tbody>
</table>

### 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
### WP4 Machine Reading System Development

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>D4.1</td>
<td>M12</td>
<td>Storing format in Knowledge Base</td>
</tr>
<tr>
<td>D4.2</td>
<td>M24</td>
<td>Multilingual platform with beta-version</td>
</tr>
<tr>
<td>D4.3</td>
<td>M24</td>
<td>SRL Software release</td>
</tr>
<tr>
<td>D4.4</td>
<td>M36</td>
<td>Multilingual platform with final version</td>
</tr>
<tr>
<td>D4.5</td>
<td>M36</td>
<td>Taxonomy induction software release</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP4 Machine Reading System Development</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>4.1 Development of Multilingual platform</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>4.2 Development of French Reading Machine</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>4.3 Development of English Reading Machine</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>4.4 Software release</td>
<td>ED</td>
<td></td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
<th>Task leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>D5.0</td>
<td>M6</td>
<td>New definition of WP5</td>
<td>SD</td>
</tr>
<tr>
<td>D5.1</td>
<td>M12</td>
<td>Report on building comprehensive business case</td>
<td>SD</td>
</tr>
<tr>
<td>D5.2</td>
<td>M24</td>
<td>Report on the design of the demonstrator and the use case</td>
<td>SD</td>
</tr>
<tr>
<td>D5.3</td>
<td>M32</td>
<td>Demonstrator Release</td>
<td>SD</td>
</tr>
</tbody>
</table>
## WP6 Evaluation of Machine Reading

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Month of delivery</th>
<th>Title of deliverable</th>
</tr>
</thead>
<tbody>
<tr>
<td>D6.1</td>
<td>M12</td>
<td>1&lt;sup&gt;st&lt;/sup&gt; Report on the lab-based evaluation</td>
</tr>
<tr>
<td>D6.2</td>
<td>M24</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt; Report on the lab-based evaluation</td>
</tr>
<tr>
<td>D6.3</td>
<td>M36</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt; Report on the lab-based evaluation</td>
</tr>
</tbody>
</table>

### Task leader
- UNED

<table>
<thead>
<tr>
<th>WP6 Evaluation of Machine Reading</th>
<th>Task leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Definition of the campaign exercises and goals</td>
<td>UNED</td>
</tr>
<tr>
<td>6.2 Evaluation methodology and resources</td>
<td>UNED</td>
</tr>
<tr>
<td>6.3 Activation of the campaign exercises</td>
<td>UNED</td>
</tr>
<tr>
<td>6.4 Assessment and comparison of systems results</td>
<td>UNED</td>
</tr>
<tr>
<td>6.5 Workshop</td>
<td>UNED</td>
</tr>
</tbody>
</table>
# Milestones

<table>
<thead>
<tr>
<th>No of Milestone</th>
<th>Delivery month</th>
<th>WP involved</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>M12</td>
<td>WP2</td>
<td>Initial Proposition Store</td>
</tr>
<tr>
<td>M2</td>
<td>M24</td>
<td>WP2, WP4</td>
<td>Semantic Role Labelling Software release</td>
</tr>
<tr>
<td>M3</td>
<td>M36</td>
<td>WP3, WP4, WP6</td>
<td>Machine Reading system</td>
</tr>
<tr>
<td>M4</td>
<td>M36</td>
<td>WP4, WP5</td>
<td>Industrial demonstrator release</td>
</tr>
</tbody>
</table>
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)

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

<table>
<thead>
<tr>
<th>N°</th>
<th>Partner</th>
<th>Person.months</th>
<th>Total costs (Euros)</th>
<th>Percentage of requested budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UNED</td>
<td>6</td>
<td>0</td>
<td>0 %</td>
</tr>
<tr>
<td>2</td>
<td>U. Edinburgh</td>
<td>12</td>
<td>127,000</td>
<td>33 %</td>
</tr>
<tr>
<td>3</td>
<td>Synapse</td>
<td>22</td>
<td>130,882</td>
<td>25.36 %</td>
</tr>
<tr>
<td>4</td>
<td>UPV/EHU</td>
<td>6</td>
<td>0</td>
<td>0 %</td>
</tr>
</tbody>
</table>

Note:
- UNED and UPV/EHU received funds at the end of the period (December 2013)
- Personnel budget only for new contracts
Website: nlp.uned.es/readers-project/

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.
1. Initial Proposition Store
   – Used to automatically unfold transfers of meaning in eventive propositions (CICLING 2014)


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


Single partner international conferences


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
  – Continuously 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