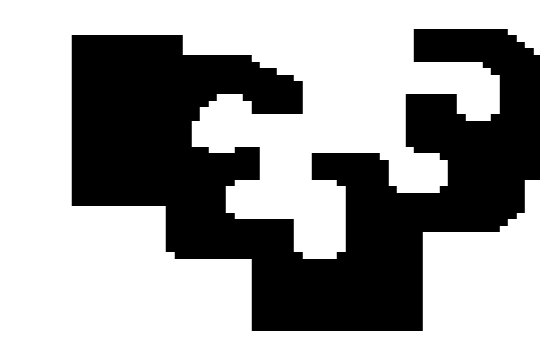




<http://ixa.si.ehu.es/>

University of the Basque Country (UPV/EHU)



UPV/EHU



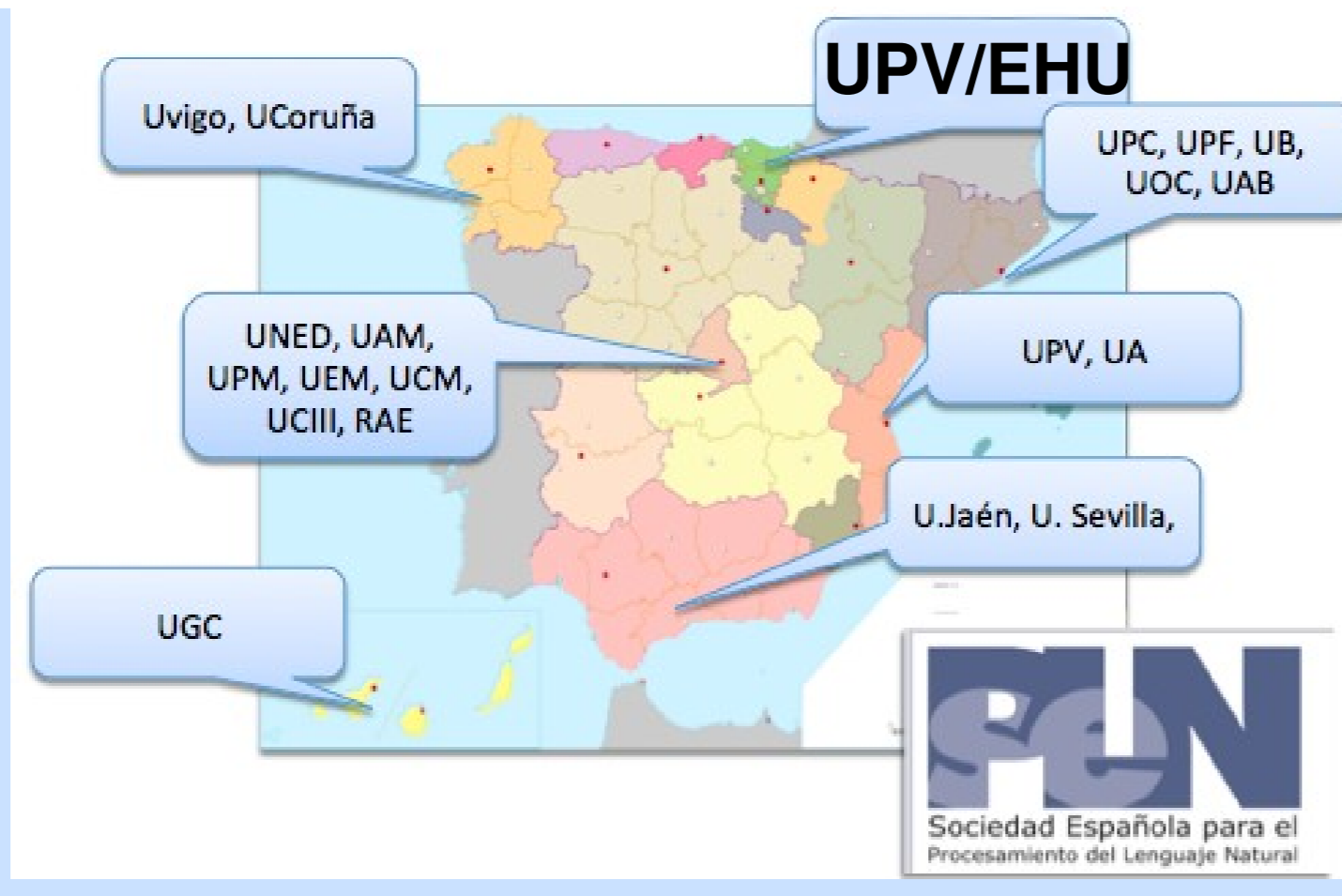
Who are we?

25+ anniversary

Interdisciplinary team

- 38 CS and 16 Linguists (38 PhDs)
- 30 teachers, 10 post-docs, 14 PhD students
- 3 research assistants

Other NLP groups in Spain



Publications and Projects

Around 50 research papers / year
Active research projects:

- 2 EU projects
- 4 National projects (Mineco)
- 3 Basque government

Main focus:

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



Our current major projects are the following: PATHS, NewsReader, READERS, OpeNER, European STREP projects; Basque Government's Stablished Research Group: Hibrido Sint, OpenMT-2, KNOW2, TACARDI and SKaTeR projects from Spanish Ministry of Education and Science; and also the EUROPEANA1, Flarenet2, TIMM and RTTH research networks.

Click [here](#) for the whole list of projects since 1988.

Read more

PROJECTS



Current research on NLP and Education

Generation of Didactic Resources (DR)

Generation of DRs from corpora: Traditional corpora and Collaboratively Constructed Resources

- Tests automatically generated
- Aim: the automatic generation of MCQs to assess learners in different environments: language learning and domain specific learning:
 - Identification of the stems
 - Automatic generation of the distractors

Generation of DRs based on existing didactic material

- Processing of Data Collection using NLP techniques to detect and extract document fragments corresponding to didactic resources (DR)
- DRs of different granularity, such as, definitions, examples, exercises, theorems, etc...
- Definition of a DR Grammar

Collaboration with the Galan research group

Giving feedback to the students

Comparing reference answers against students responses

Techniques: Semantic Textual Similarity (score) and Natural Language Inference (Label)

Datasets available at Semeval workshop

- Systems:**
- Machine Learning based
 - Deep learning based

Networks

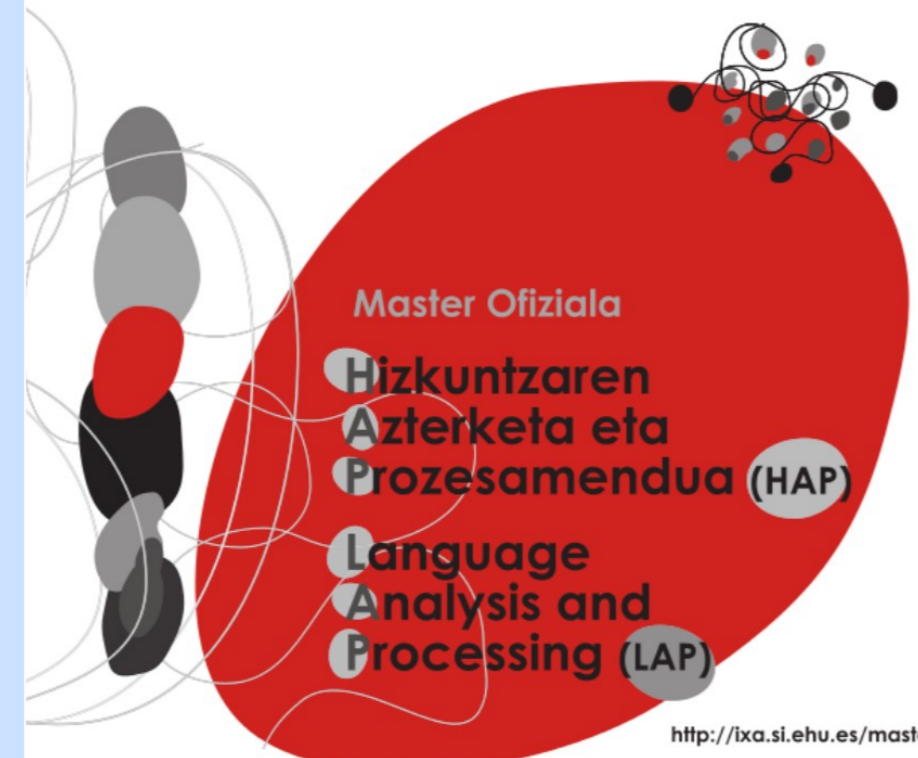
- Speech research team (Aholab)
- Corpus monitor for the **Basque Academy**
- Lexicography (Elhuyar): most popular dictionary
- Spin-off company (Eleka)
- Universities, MetaNet...

Education

- Master**
- HAP/LAP Language Analysis and Processing

- Master Erasmus-Mundus**
- Language Technology and Communication (2013-2018)
<http://lct-master.org/>
 - Saarbrücken (coordinator), Groningen, Trento, Lorraine, Malta, Prague, Melbourne, Shanghai and UPV/EHU

- PhD program:**
- HAP Language Analysis and Processing
 - 15 PhDs from 2010



Other research interests

- Playing with literary texts
- Generation of who-questions
- Reading Comprehension
- Essay scoring
- Building Educational Datasets
- ...