INTRODUCTION TO EVALITA 2009

Bernardo Magnini[1] and Amedeo Cappelli[2,3]

[1]Fondazione Bruno Kessler
[2] CELCT, Center for the Evaluation of Language and Communication Technologies
[3] ISTI, Institute of Information Science and Technologies, Italian National Research Council

In its second edition, EVALITA has become the reference evaluation campaign of both Natural Language Processing and Speech technologies for Italian. The objective of the shared tasks proposed at EVALITA is to promote the development of language technologies for the Italian language, providing a common framework where different systems and approaches can be evaluated and compared in a consistent manner.

Thanks to the support of the Italian Association of Speech Science (AISV), the major novelty of EVALITA 2009 is that, for the first time, an evaluation initiative of Speech technologies for Italian has been organized. As a second relevant novelty, EVALITA started a collaboration with Passage (http://atoll.inria.fr/passage/index.en.html), the evaluation campaign on parsing for French, aiming at sharing methodologies with similar initiatives at the European level.

Both academic institutions and industrial organizations have been invited to participate at EVALITA 2009, and 102 expressions of interest have been received for the eight tasks proposed. In the end, the number of participants was 46, with the following distribution over the tasks:

Text tasks
- PoS-Tagging: 8
- Parsing: 8
- Lexical Substitution: 3
- Entity Recognition: 8
- Textual Entailment: 2

Speech tasks
- Connected Digits Recognition: 4
- Spoken Dialogue Systems Evaluation: 3
- Speaker Identity Verification: 10

Overall, 30 different organizations participated in EVALITA 2009, out of which 11 were not Italian, and 5 were non-academic, confirming the interest in the proposed tasks.

As with other international evaluation campaigns, participants have been provided with training data and had the chance to test their systems with the evaluation metrics and procedures to be used in the formal evaluation. As for EVALITA 2007, after the campaign, most of the datasets are made freely available for research purposes, this way encouraging further improvement of the technologies.

We believe that the diffusion of shared tasks and shared evaluation practices is a crucial step towards the development of resources and tools for Natural Language Processing and Speech recognition. The good response obtained by EVALITA, both in the number of participants and in the quality of results, showed that it is worth pursuing such goals for the Italian language.

We would like to thank all the people and the institutions involved in the organization of the EVALITA tasks. Overall, it has been a tremendous effort, and we are really proud that so many persons and organizations were able to achieve a so ambitious result on a completely voluntary basis.

We are grateful to the interest group on NLP of the AI*IA, who encouraged this initiative since its beginning in 2007 and hosted the final workshop in Reggio Emilia, and to AISV, both for the contribution in the organization of the speech task and for the sponsorship. The collaboration with the French initiative Passage has been made possible thanks to the cooperation with Patrick Paroubek, from LIMSI, who kindly accepted to give an invited talk at the workshop. We also thank the LiveMemories project (www.livememories.org) and the International Master in Human Language Technologies and Interfaces at the University of Trento, for their kind sponsorships.

Finally, a special thank to Manuela Speranza, who took care of most of the managing work at FBK, and to Valentina Bartalesi Lenzi and Rachele Sprugnoli, who managed the EVALITA Web site and the realization of the proceedings at CELCT.