



EVALITA 2011

Evaluation of NLP and Speech Tools for Italian

EVALITA 2011

Third workshop on Evaluation of Natural Language Processing and Speech Tools for Italian

Bernardo Magnini – Emanuele Pianta

Rome, January 24-25 2012



EVALITA: Objectives

- Promote the development and diffusion of NLP and Speech tools for Italian
- Provide shared tasks
 - Training and test data
 - Common evaluation procedures
- Take advantage of existing annotated corpora for Italian
- Diffusion/re-use of datasets



NLP Tasks at EVALITA 2011

1. Lemmatization (Uni. Bologna)
2. Parsing (Uni. Torino)
3. Domain adaptation (CNR-ILC, Pisa)
4. Entity recognition on transcribed news (FBK, Celct, Trento)
5. Cross document coreference (FBK-irst, CELCT, Trento)
6. Anaphora resolution (Uni. Trento)
7. Super sense tagging (Uni. Pisa)
8. Frame labeling (Uni. Tor Vergata, Uni. Pisa)

5 new tasks, 2 with increased complexity, 1 without changes



EVALITA NLP: Participation

	Participation
Lemmatization	4 (5)
Parsing (Dependency + Constituency)	5 (18)
Domain adaptation	2 (7)
Entity Recognition on transcribed broadcast news	4 (14)
Cross document coreference resolution	1 (6)
Anaphora resolution	1 (4)
Super sense tagging	2 (7)
Frame labeling over Italian	3 (8)

8 NLP tasks, 69 registered, 22 actual participants

- 16 different organizations
- 3 non Italian groups, 2 non academic



EVALITA 2011: Discussion

- Performance: for parsing results have constantly improved over the years
- Consolidated tasks (e.g. well known format) are preferred
- Datasets for new tasks are now available: we need to capitalize on the big effort
 - reuse datasets for benchmarking: post-Evalita publicity
 - publications
 - this year we have post-proceeding edited by Springer
 - special issue of "Intelligenza Artificiale", IOS Press
 - International visibility of the Italian community: position Evalita datasets within EU initiatives



EVALITA: Sponsors



Italian Association for Speech Sciences



Italian Association for Artificial Intelligence



Workshop Program

- More time: one day and a half!
- Task sessions
 - Tasks and results presentation
 - Selected presentations
 - Discussion
- Poster Session (and coffee break), this morning
- Presentation of the Parli portal (this afternoon)
- Panel on Bridging NLP and Speech (tomorrow morning)
- Student context (tomorrow morning)



Thank You!

- Evalita has been organized on a voluntary basis:
 - Many thanks to task organizers!
 - Many thanks to all participants!
 - Many thanks to AISV and local organizers!



Thank You!

- Evalita has been organized on a voluntary basis:
 - Many thanks to task organizers!
 - Many thanks to all participants!
 - Many thanks to AISV and local organizers!

Special thank to:

Manuela Speranza (FBK-irst), Valentina Bartalesi Lenzi (CELCT), Rachele Sprugnoli (CELCT)