



EVALITA 2011

Evaluation of NLP and Speech Tools for Italian

EVALITA 2011

Named Entity Recognition on Transcribed Broadcast News Task

V. Bartalesi Lenzi, M. Speranza, R. Sprugnoli





Introduction

- **Aim:** Recognize Named Entities in Italian texts
- **4 types** of Named Entities (ACE formalism adapted to Italian):
 - Person Entities (PER): e.g. *Napolitano*
 - Organization Entities (ORG): e.g. *FIAT*
 - Geo-Political Entities (GPE): e.g. *Italia*
 - Location Entities (LOC): e.g. *Tevere*
- **Novelties:**
 - task based on automatic transcription of news broadcasts:
 - transcription errors, no sentence boundaries, no punctuation
 - 2 subtasks:
 - 1) Full: automatic transcription + NE recognition
 - 2) NER only: NE recognition on the automatic transcript provided by the organizers



Participants and Runs

- 14 groups registered to the task but only 4 submitted the results all for the *NER only* subtask :
 - FBK_Alam
 - FBK_Chowdhury
 - UniPi_SimiDeiRossi
 - ISI_GhoshKozareva
- 2 types of run: *closed modality* and *open modality*
- Total of 9 runs:
 - 6 produced in a closed modality
 - 3 produced in an open modality



Dataset

- 20 broadcast news provided by RTTR

	Training	Test
Broadcast news	10	10
Hours of transmission	5	5
# Tokens	42,595	36,643
# Named Entities	1,924	1,703

- **Training Data:**

- manual transcriptions annotated with NEs
- automatic transcriptions produced using FBK's ASR system
- possibility to use I-CAB

- **Test Data:**

- automatic transcriptions produced using FBK's ASR system



Evaluation Procedure

- **Gold standard:** manual transcriptions annotated with NEs by three expert annotators

- **3 phases:**
 - 1) **transcription alignment:** align the gold standard with the system's transcription
 - 2) **NER error detection:** compare the NEs annotated in the gold standard to the NEs annotated by NER systems in the automatic transcription
 - 3) **score computation:** compute Precision, Recall and F1 –Measure with the CONLL scorer



Results

Participant	Over. FB1	FB1				
		GPE	LOC	ORG	PER	
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
- BASELINE	44.93	69.00	36.49	43.37	18.10	
- BASELINE-u	31.11	40.12	22.56	42.05	18.25	



Results

	Participant	Over.	FB1			
		FB1	GPE	LOC	ORG	PER
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
-	BASELINE	44.93	69.00	36.49	43.37	18.10
-	BASELINE-u	31.11	40.12	22.56	42.05	18.25



Results

	Participant	Over.	FB1			
		FB1	GPE	LOC	ORG	PER
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
-	BASELINE	44.93	69.00	36.49	43.37	18.10
-	BASELINE-u	31.11	40.12	22.56	42.05	18.25



Results

Participant	Over. FB1	FB1				
		GPE	LOC	ORG	PER	
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
- BASELINE	44.93	69.00	36.49	43.37	18.10	
- BASELINE-u	31.11	40.12	22.56	42.05	18.25	



Results

Participant	Over. FB1	FB1				
		GPE	LOC	ORG	PER	
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
- BASELINE	44.93	69.00	36.49	43.37	18.10	
- BASELINE-u	31.11	40.12	22.56	42.05	18.25	



Results

	Participant	Over.	FB1			
		FB1	GPE	LOC	ORG	PER
CLOSED	1 FBK_Alam_rc1	60.98	80.12	55.21	46.82	50.96
	2 FBK_Alam_rc2	60.67	78.89	56.25	47.77	49.36
	3 FBK_Chowdhury_rc	57.02	74.61	51.01	42.14	46.10
	4 UniPi_SimiDeiRossi_rc2	56.61	76.18	44.30	38.35	46.81
	5 UniPi_SimiDeiRossi_rc1	53.47	73.60	43.97	28.79	46.77
	6 ISI_GhoshKozareva_rc	42.42	64.21	32.52	15.01	28.74
OPEN	1 FBK_Alam_ro1	63.56	80.38	56.38	53.24	51.51
	2 UniPi_SimiDeiRossi_ro1	58.19	76.25	48.78	40.60	48.75
	3 UniPi_SimiDeiRossi_ro2	52.15	72.25	30.62	33.77	46.19
-	BASELINE	44.93	69.00	36.49	43.37	18.10
-	BASELINE-u	31.11	40.12	22.56	42.05	18.25



Conclusions

- New task: decrease in the number of participants (4 groups)
 - 6 groups in 2007 and 7 in 2009
- No submissions to the *Full Subtask*
- Decrease in systems performance is significant with respect to written news stories (see EVALITA 2009 and 2007)
 - decrease of 20 points for the best system
 - in line with the WER of the ASR system
- Future work: evaluate the impact of transcription errors on NER performance



THANK YOU!

Special thanks to Diego Giuliani and Roberto Gretter for their helpful collaboration and advice as ASR experts