

# **An Ontology-driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations**

—Appendix—

Marco Rospocher

Fondazione Bruno Kessler – IRST  
Via Sommarive 18, Trento, I-38123, Italy  
rospocher@fbk.eu

COMPLEMENTARY TABLES WITH THE SCORES FOR ALL METRICS AND  
EVALUATION MEASURES CONSIDERED IN THE PAPER:

Marco Rospocher

*An Ontology-driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations*  
In Proceedings of the 17th International Semantic Web Conference, ISWC 2018  
Monterey, CA, USA, October 8-12, 2018

Table 1: Micro-averaged results (only gold standard mentions).

		type			link			type+link		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
AIDA	<i>standard</i>	.943	.875	.908	.662	.652	.656	.634	.625	.630
	<i>with PSLAEA</i>	.947	.879	.912	.670	.659	.665	.646	.635	.640
	$\Delta$	.004	.004	.004	.008	.007	.009	.012	.010	.010
MEANTIME	<i>standard</i>	.882	.695	.777	.703	.556	.621	.635	.502	.561
	<i>with PSLAEA</i>	.902	.711	.795	.714	.564	.630	.667	.527	.589
	$\Delta$	.020	.016	.018	.011	.008	.009	.032	.025	.028
TAC-KBP	<i>standard</i>	.911	.652	.760	.401	.423	.412	.367	.386	.376
	<i>with PSLAEA</i>	.925	.662	.772	.408	.430	.419	.384	.404	.394
	$\Delta$	.014	.010	.012	.007	.007	.007	.017	.018	.018

Table 2: Micro-averaged results (all mentions). Note: Precision and F<sub>1</sub> scores on TAC-KBP are extremely low (and ignored in the summary reported in the paper) due to the fact that only mentions of one single entity per document is annotated in the gold standard (4969 mentions), while the tools annotate all the content of each document (116009 mentions).

		type			link			type+link		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
AIDA	<i>standard</i>	.881	.875	.878	.634	.652	.642	.608	.625	.616
	<i>with PSLAEA</i>	.884	.879	.881	.642	.659	.651	.619	.635	.627
	$\Delta$	.003	.004	.003	.008	.007	.009	.011	.010	.011
MEANTIME	<i>standard</i>	.484	.695	.570	.430	.556	.485	.388	.502	.438
	<i>with PSLAEA</i>	.495	.711	.584	.436	.564	.492	.408	.527	.460
	$\Delta$	.011	.016	.014	.006	.008	.007	.020	.025	.022
TAC-KBP	<i>standard</i>	.028	.652	.054	.012	.423	.023	.011	.386	.021
	<i>with PSLAEA</i>	.028	.662	.054	.012	.430	.024	.011	.404	.022
	$\Delta$	.000	.010	.000	.000	.007	.001	.000	.018	.001

Table 3: Macro-averaged (by Document) results (using only gold standard mentions).

		type			link			type+link		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
AIDA	<i>standard</i>	.939	.861	.895	.646	.651	.636	.616	.622	.607
	<i>with PSLAEA</i>	.942	.864	.898	.663	.671	.654	.633	.642	.625
	$\Delta$	.003	.003	.003	.017	.020	.018	.017	.020	.018
MEANTIME	<i>standard</i>	.886	.718	.776	.684	.551	.595	.630	.509	.548
	<i>with PSLAEA</i>	.905	.732	.792	.693	.560	.604	.659	.533	.574
	$\Delta$	.019	.014	.016	.009	.009	.009	.029	.024	.026
TAC-KBP	<i>standard</i>	.811	.695	.733	.260	.231	.241	.233	.207	.216
	<i>with PSLAEA</i>	.825	.707	.746	.265	.236	.246	.246	.220	.228
	$\Delta$	.014	.012	.013	.005	.005	.005	.013	.013	.012

Table 4: Macro-averaged (by NERC Type) results (using only gold standard mentions).

		type			link			type+link		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
AIDA	<i>standard</i>	.937	.863	.898	.601	.584	.592	.601	.584	.592
	<i>with PSLAEA</i>	.942	.866	.902	.615	.596	.604	.615	.596	.604
	$\Delta$	.005	.003	.004	.014	.012	.012	.014	.012	.012
MEANTIME	<i>standard</i>	.960	.680	.795	.606	.437	.508	.606	.437	.508
	<i>with PSLAEA</i>	.978	.693	.810	.626	.454	.525	.626	.454	.525
	$\Delta$	.018	.013	.015	.020	.017	.017	.020	.017	.017
TAC-KBP	<i>standard</i>	.924	.666	.772	.368	.397	.381	.368	.397	.381
	<i>with PSLAEA</i>	.936	.678	.783	.382	.415	.397	.382	.415	.397
	$\Delta$	.012	.012	.011	.014	.018	.016	.014	.018	.016