



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

Working Group 4 - Report

Irene Russo - ILC CNR (WG leader)

Liviu P. Dinu - University of Bucharest (WG co-leader)



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 main aims

Exploiting large knowledge bases and graphs

- This WG focuses on the exploitation of diverse structured KBs and language resources for multimodal LG tasks. The analysis of how to efficiently integrate or enrich these (multimodal) KBs is paired with the study of theoretical models of semantics and semantic processing that can accommodate linguistic and perceptual information



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 tangible outcomes

- 1) In-depth analysis and review of methods for incorporating structured knowledge bases into language generation models
- 2) Literature review and state-of-the-art on linguistic theories combining linguistic and perceptual information in semantic representations



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 deliverables

- an email was sent after December 2020 meeting, asking for collaboration on deliverables
- two researchers were interested in collaborating on D2
- no one reacted for D1
- the deliverables have been structured and missing content could be easily provided
- I will ask for comments/integrations



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 - [D1 structure](#)

- Introduction
- Knowledge base and knowledge graph
- General-purpose knowledge bases (with papers that use them)
- Data-to-text NLG training datasets
- A taxonomy of NLG systems with respect to knowledge bases



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 - D2 structure

- Introduction
- Theories of semantic meaning representation and conceptual combination
- Insights from theories
- Possible uses of psycholinguistic data in NLG
 - list of behavioural datasets (association norms, semantic priming, word substitution errors etc.)
- Conclusions



MULTI-TASK
MULTI-LINGUAL
MULTI-MODAL



Funded by the Horizon 2020 Framework Programme
of the European Union

WG4 general outputs

Publications from deliverables?

- concerning D1, there is a recent survey about data-to-text NLG
- Yu et al. (2020), A Survey of Knowledge-Enhanced Text Generation
- it focuses on methodologies and less on resources
- publish on the website a lists of KBs with papers that use them?

- concerning D2, there are a lot of surveys about semantic representations in psycholinguistics and cognitive psychology
- proof-of-concept paper on deep learning LMs and psycholinguistic datasets? -> Cross-topic research roadmap
- survey on creative NLG?

NooJ introductory course for a summer school?

- Anabela got me in touch with Max Silberztein