Individual APT Proposal - 2022/2023

Enter the details related to the GAPT that you want to propose below. This year we should try as much as possible to propose tasks that can be performed by groups of three students.

Title *

Triples2Text

Supervisor *

Claudia Borg

Co-Supervisor

Main Subject Area *

Natural Language Processing

Short Description of the Task *

This task consists in mapping data to text. The training data consists of Data/Text pairs where the data is a set of triples extracted from DBpedia and the text is a verbalisation of these triples. For instance, given the 3 DBpedia triples shown in (a), the aim is to generate a text such as (b).

a. (John_E_Blaha birthDate 1942_08_26) (John_E_Blaha birthPlace San_Antonio) (John_E_Blaha occupation Fighter_pilot)
b. John E Blaha, born in San Antonio on 1942-08-26, worked as a fighter pilot

As the example illustrates, the task involves specific NLG subtasks such as sentence segmentation (how to chunk the input data into sentences), lexicalisation (of the DBpedia properties), aggregation (how to avoid repetitions) and surface realisation (how to build a syntactically correct and natural sounding text).

Deliverables *

Code, report

Recommended Resources (including readings) *

https://synalp.gitlabpages.inria.fr/webnlg-challenge/challenge_2020/
https://aclanthology.org/2020.webnlg-1.6.pdf
https://aclanthology.org/2020.webnlg-1.7.pdf

Prerequisite Knowledge Required

Currently taking ICS2203

List any Ethical issues

This form was created inside University of Malta.

Google Forms