

## EUROPROtweets Labelling and Tagging

The verbal component of the tweets from 20 research project websites linked to a Twitter account (out of the 30 websites selected for EUROPROwebs pilot corpus) was downloaded and saved into a txt-format document corresponding to each account labelled NAMEOFTHEPROJECT\_T.

The tagging was carried out as follows:

1. **Manual tagging:** A number of tags was determined in the light of the texts downloaded and saved. Two main foci were of interest to us:

a) **Interactive affordances:** <user>, <hashtag>, <mention>, <hyperlink>

b) **Multimodal elements:** <image>, <emoji>, <video>, <gif>, <meme>

The whole corpus was thus annotated.

Hammerschmid, <mention>@kaiwegrich1</mention> <mention>@EU\_H2020 project</mention> “TROPICO” to research <hashtag>#egovernance</hashtag>, <hashtag>#digitalisation</hashtag> in public sector. <link>https://t.co/fjx6F94B68</link><ext> <link>https://t.co/Yy5kEilsbL</link><ext>

Do formal rules limit or foster <hashtag>#collaboration</hashtag> in and by governments? Insights on the alleged “maze of rules” and <hashtag>#DigitalTransformation</hashtag> in public sectors from 10 European countries available now: <link>https://t.co/6XR4oRGcn3</link><int>

### Example of manual tagging in EUROPROtweets Corpus

The different types of tweets were discerned in the corpus, too, by referring to their nature and/or their form. The labels that we used are:

<TW>	To indicate user-generated tweets published in the accounts of the international research projects.
<RT>	To signal tweets written by other users and appropriated by international research projects. Quoted tweets were included within this category.
<REPLY>	To underline that the tweet was user-generated, but appeared in relation to a previous published tweet and in the form of a bidirectional response.
<THREAD> ... </THREAD>	To mark the onset and end of chains of user-generated tweets, where the content of a subsequent tweet was clearly connected with or derived from the previous one.

2. **Semi-automatic tagging:** Depending on the research questions addressed in our studies and the methods and procedures undertaken, different tagging systems were devised at the lexico-grammatical, discursive or pragmatic levels. As with EUROPROwebs Corpus, we use the software for qualitative analysis NVivo12.

By way of illustration, the figure below displays an example of a pragmatic analysis of specific strategies that are grouped around three overriding scopes: informative (INF), promotional (PRO) and interactional (INT):



Example of semi-automatic tagging in EUROPROtweets Corpus for a pragmatic study

### Related publications:

Pascual, D., Mur-Dueñas, P. and Lorés, R. 2020. Looking into international research groups' digital discursive practices: Criteria and methodological steps in the compilation of the EUROPRO digital corpus. *Research in Corpus Linguistics* 8 (2), 87-102.  
<https://doi.org/10.32714/ricl.08.02.05>

To visit the EUROPRO digital Database click [HERE](#)

