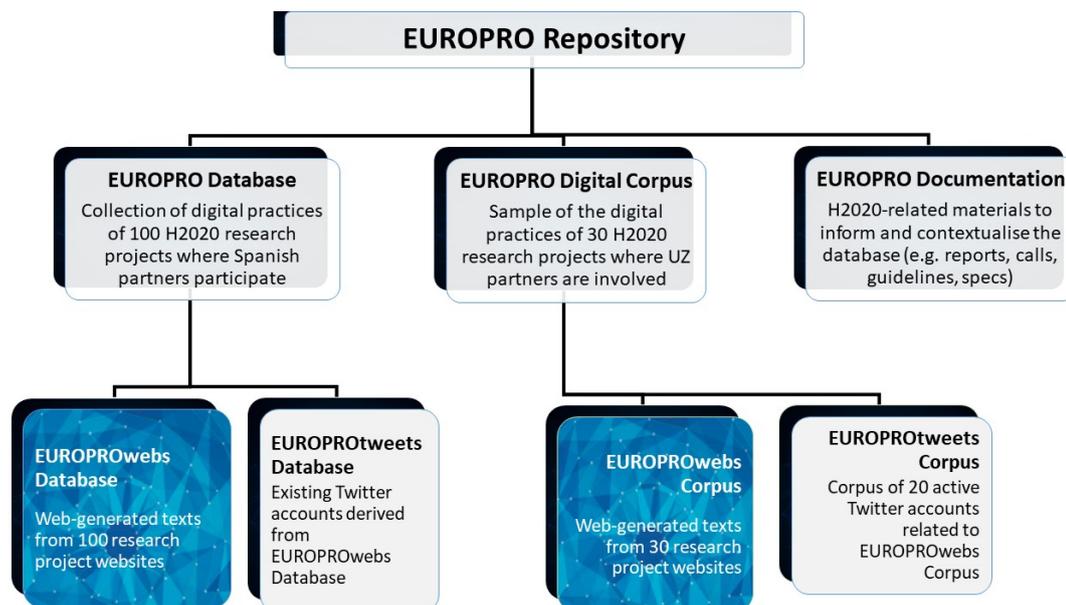


EUROPRowebs Corpus Criteria

EUROPRowebs (European Project Websites) is a pilot corpus, which contains 30 research project websites created and maintained by international projects funded under the H2020 programme. It is an ad-hoc, specialized, static corpus that represents the digital, discursive practices of international researchers collaborating in large-scale international research projects. The corpus is part of our EUROPRowebs Database comprising the research project websites of a total of 100 international projects.



Several criteria were followed for the selection of the 30 international funded projects and their websites.

1. The research projects should aim at producing new knowledge and not at training PhD students or professionals to undertake the research. This was done to ensure comparability within the corpus.
2. A convenience sampling method was followed, which entailed choosing research projects with at least a member from the Universidad de Zaragoza (Spain) or from a Zaragoza-based institution. This would allow us to complement our text-based analysis with contextual evidence from potential informants.
3. The date of the H2020 projects should at some point overlap with the development of our own project (2017-2021), so that the most recent digital academic practices could be studied.

4. Given their organic nature, the websites were selected and downloaded on a fixed date, which would entail that the duration of each of the projects was different at the moment of compilation. Since creating a monitor corpus (a corpus that is regularly updated) lies beyond the scope of our research project, websites had been developed to different stages when downloaded.

Likewise, EUROPROwebs pilot corpus was designed taking into consideration some important methodological decisions about the digital nature of the texts and their particular characteristics:

1. **DYNAMICITY** The corpus was compiled from April to May 2019. The specific start and end date of each project was retrieved in order to determine the point of development of the corresponding research project at the moment of the corpus compilation.
2. **HYPERMODALITY AND HYPERMEDIALITY.** The corpus was tagged for hyperlinks (external, internal and peripheral), for visuals (such as tables, figures, pictures, logos, etc.), for videos and audios.
3. **LAYOUT AND WEB DESIGN.** Besides downloading the texts contained in the website, screenshots of every page were taken and saved. Information was also recorded of the extent to which the text could be directly accessed from the website menu.

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](#)

