

At Least 21 European Languages in Danger of Digital Extinction

Study by Europe's Leading Language Technology Experts Warns Most European Languages Unlikely to Survive in the Digital Age

Most European languages face digital extinction, a new study by Europe's leading Language Technology experts finds. Assessing the level of support through language technology for 30 of the approximately 80 European languages, the experts conclude that digital support for 21 of the 30 languages investigated is "non-existent" or "weak" at best. The study was carried out by META-NET, a European network of excellence that consists of 60 research centres in 34 countries.

The study, prepared by more than 200 experts and documented in 30 volumes of the META-NET White Paper Series (available both online and in print), assessed language technology support for each language in four different areas: automatic translation, speech interaction, text analysis and the availability of language resources. A total of 21 of the 30 languages (70%) were placed in the lowest category, "support is weak or non-existent" for at least one area by the experts. Several languages, for example, Icelandic, Latvian, Lithuanian and Maltese, receive this lowest score in at least three areas. On the other end of the spectrum, while no language was considered to have "excellent support", only English was assessed as having "good support", followed by languages such as Dutch, French, German, Italian and Spanish with "moderate support". Languages such as Basque, Bulgarian, Catalan, Greek, Hungarian and Polish exhibit "fragmentary support", placing them also in the set of high-risk languages.

Europe has succeeded in removing almost all borders between its countries. But the invisible borders created by language barriers continue to obstruct the exchange of knowledge and information and hence, the free flow of goods and services, thus harming the long-term goal of establishing a single digital market. Even though language technology has the potential to lower or even eliminate language barriers through machine translation systems and related tools, the results of the META-NET study clearly show that many of European languages are not yet ready. There are significant gaps in technology due to the English-language focus of most R&D, a lack of commitment and financial resources, and also a lack of a clear research and technology vision.

Europe must take action to prepare its languages for the digital age. They are a precious component of our cultural heritage and, as such, they deserve future-proofing. On September 26th the Council of Europe marks the European Day of Languages, a day which recognises the importance of fostering and developing the rich linguistic and cultural heritage of our continent. META-NET's work is a stark reminder of the challenges and possibilities facing our linguistic heritage in the information age.