

Language Technology News – Iceland

Eiríkur Rögnvaldsson
ELG NCC Iceland

META-FORUM 2019
Brussel, October 9



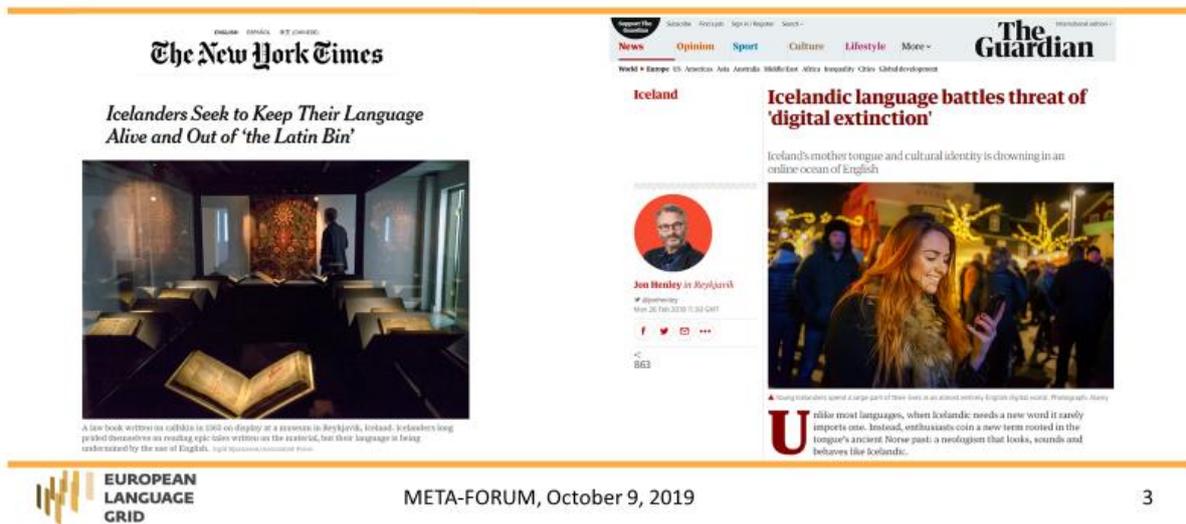
THE ÁRNI MAGNÚSSON INSTITUTE
FOR ICELANDIC STUDIES

The META-NET White Paper



Even though language technology work in Iceland started around the turn of the century, I think it is fair to say that Iceland’s participation in META-NET was a landmark in Icelandic language technology. The White Paper on Icelandic highlighted the alarming status of Icelandic LT. Icelandic was one of only four languages receiving the lowest score in all four categories that were evaluated – which is perhaps not surprising, given that Icelandic only has 350,000 speakers.

Icelandic in international media



Following the publication of the White Paper, we started extensive lobbying and propaganda for Icelandic LT. In recent years, the prospective future of the Icelandic language has become the subject of much discussion in Iceland and even internationally. Influential newspapers have recently published articles on Icelandic in the digital age, emphasizing the threats that the digital revolution poses to Icelandic and other small languages.

Contacts with international IT companies



The President of Iceland visiting Microsoft headquarters



The President testing Icelandic on the Microsoft Translator app

We have made efforts to establish contacts with the big international IT companies in order to try to persuade them to include Icelandic in their products. Last year, the President of Iceland visited Microsoft headquarters in Seattle, accompanied by representatives from Language technology, both from academia and industry. On that occasion, Microsoft announced that Icelandic had been added to Microsoft Translator which the President is testing here.

The White Paper discussed in the Parliament



However, the campaign has of course mainly been fought within Iceland. The alarming results of the White Paper were discussed in the Icelandic Parliament in 2012 as you see on this picture. In 2014, the Parliament unanimously adopted a resolution on the necessity of making Icelandic usable in the digital domain. In 2016, the Government commissioned three language technology experts to develop a detailed five-year Project Plan for Icelandic language technology.

Language Technology Project Plan, 2018-2022

- Five core tasks
 - Speech recognition
 - Speech synthesis
 - Machine translation
 - Language & style checking
 - Language resources
- Competitive R&D Fund
 - Financed partly (50%) by the industry
- LT Education
- CLARIN membership



- Costs
 - Core tasks
 - 9,920,000 €
 - Competitive R&D Fund
 - 7,430,000 €
 - LT Education
 - 740,000 €
 - Participation in CLARIN
 - 710,000 €
 - LT Centre
 - 1,410,000 €
- Total costs
 - 20,210,000 €

The Project Plan which was delivered in June 2017 proposes that emphasis be laid on five core tasks. Other proposals include establishing a competitive research and development LT fund, strengthening LT education, and joining CLARIN ERIC. In the policy statement of the current Government, it is explicitly stated that the Project Plan will be implemented and financed. The estimated total cost of the project is around 20.2 million Euros.

Signing of contracts on the LT Project Plan



The Minister of Education, Science and Culture signing contract with Almannarómur, August 2018

Almannarómur signing contract with the SÍM Consortium for Language Technology, September 2019



In August 2018, the Government commissioned the self-owned foundation Almannarómur to conduct the five above-mentioned core tasks. Last month, a contract was signed between Almannarómur and the SÍM Consortium that will carry out the research and development work necessary for these tasks. The deliverables will be stored and distributed under open licenses by the Icelandic CLARIN Centre – and I presume they could also become part of the ELG.

The SÍM Consortium for Language Technology

	UNIVERSITY OF ICELAND	University of Iceland (UI)
	HÁSKÓLINN Í REYKJAVÍK REYKJAVÍK UNIVERSITY	Reykjavik University (RU)
	The Árni Magnússon Institute for Icelandic Studies	The Árni Magnússon Institute for Icelandic Studies
	RÚV	RÚV, The Icelandic National Broadcasting Service
	BLINDRAFÉLAGIÐ	Icelandic Association of the Visually Impaired
	CREDITINFO	Creditinfo Ltd.
	MÍÐEIND	Miðeind Ltd.
	TIRO	Tiro Ltd.
	GRAMMATEK	Grammatek Ltd.



The LT Project Plan execution emphasizes getting industry involved as early as possible in the process. This is reflected in the SÍM consortium which consists of three academic institutions, the National Broadcasting Service, the Association of the Visually Impaired, and four private IT companies. Thus, the consortium comprises practically all institutions, companies, and individuals that have been involved in LT in Iceland for the past two decades – and some new ones.

The SÍM Consortium has been firmly established and the consortium members are busy planning the work ahead and recruiting and training staff. Work on all the core tasks has started already or will start in the near future. The research and development fund has already given out a few relatively small grants, and some projects supported by the fund are well underway. It is expected that larger three-year grants will be advertised in the course of the next months.

A joint master's program in Language Technology which the University of Iceland and Reykjavík University established in 2007 has now been resurrected and strengthened and several students are enrolled. Furthermore, three LT PhD-students have started their studies. Following the proposal in the Project Plan, Iceland has now joined CLARIN ERIC – only as an observer, but Icelandic law has now been changed so Iceland can apply for full membership.

The Icelandic Language Technology Project Plan is very ambitious and in order for it to go through, we will need all the advice and assistance we can get. European cooperation is crucial in this respect – I already mentioned META-NET, and we are confident that participation in the European Language Grid will be very valuable for our work in the next few years. So, at the moment, the future of Icelandic Language Technology looks promising!