



Lithuania in European Language Technology Research Infrastructures

Andrius Utkas

Centre of Computational Linguistics /

CLARIN centre

Lithuania in European Language Technology Research Infrastructures

■ **International Research Infrastructures**

- META-NET - Multilingual Europe Technology Alliance (coordinator –Institute of Lithuanian language, joined in 2011),
- CLARIN - Common Language Resources and Technology Infrastructure (coordinator Vytautas Magnus university, joined in 2014),

Lithuanian resources in Meta-Share Resources

- Meta-Nord node contains 66 Lithuanian resources,
- As a user of Meta-NORD, I noticed that the number of Lithuanian resources has not changed since 2014,

What is the most popular Lithuanian resource in Meta-Share?

66 Language Resources (Page 1 of 4)

« Previous | [Next](#) »

Order by: ▼

 **ACCURAT balanced test corpus for under resourced languages**   9  39
[Croatian](#) [English](#) [Estonian](#) [German](#) [Greek](#) [Latvian](#) [Lithuanian](#) [Romanian](#) [Slovenian](#)

 **ACCURAT corpus of comparable sentences**   16  37
[Croatian](#) [English](#) [Estonian](#) [German](#) [Greek](#) [Latvian](#) [Lithuanian](#) [Romanian](#) [Slovenian](#)

 **ACCURAT corpus of Wikipedia texts**   7  37
[Croatian](#) [English](#) [Estonian](#) [German](#) [Greek](#) [Latvian](#) [Lithuanian](#) [Romanian](#) [Slovenian](#)

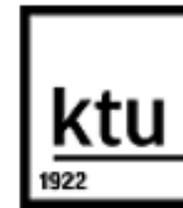
 **ACCURAT test corpus for renewable energy domain**   6  30
[Croatian](#) [English](#) [Latvian](#) [Lithuanian](#) [Romanian](#)

18 months passed from the foundation of CLARIN-LT consortium

- 25/10/2014 Lithuania joined CLARIN ERIC;
- 31/03/2015 Vytautas Magnus University, Kaunas University of Technology and Vilnius University founded the CLARIN-LT consortium.
- 25/06/2015 CLARIN-LT received funding and started activities.
- The size of the funding for two years has been 350 thousand euros (until 2016).



Lietuvos
mokslo
taryba



kaunas
university of
technology



New National Coordinator



Dr. Jurgita Vaičėnėnė

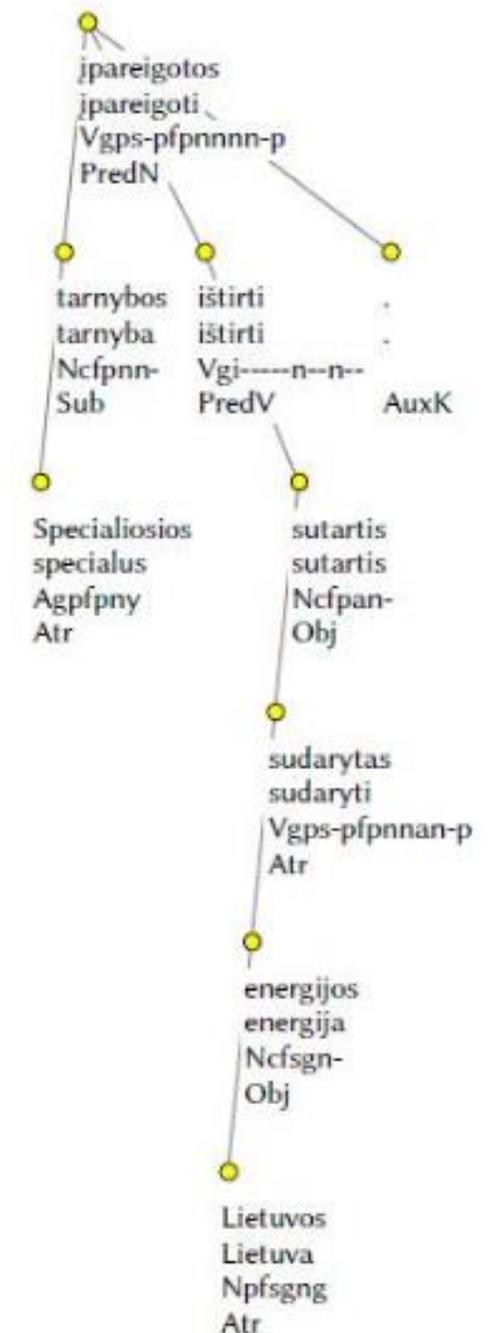


Main activities

- to connect to CLARIN ERIC network, by building a B type centre in Kaunas, Lithuania;
- to improve services and resources in the existing infrastructure;
- to consult language researchers;
- to ensure harmonisation of legal issues between CLARIN ERIC and CLARIN LT including the usage of appropriate licences;
- to advertise new possibilities for language researchers and involve new users.

Improving existing resources and creating new ones

- Building a speech corpus (15hrs);
- Increasing the existing Web Corpus (by 400M words);
- LitWordNet (20 thousand synsets);
- Creation of a syntactic Treebank ALKSNIS (2,300 syntactic trees);
- Increasing morphological and grammatical databases.
- Developing a new search engine for Contemporary Lithuanian language corpus with new functionalities (*sorting, lemma search, morphological category search*).



Future prospects

- “An inclusion of any project into the national Roadmap means recognition of its national significance for the progress and breakthrough of Lithuanian research.”
- Based on that we expect to receive further funding for activities of CLARIN-LT consortium. However, the amount of the funding remains vague.



Thank you!
Ačiū už dėmesį!