



Events: Named Entity Recognition in Digital Humanities Workshop, Helsinki

Tweet #nerhelsinki

Searching for strings that have name like properties is the most frequent type of search in information retrieval. Named Entity Recognition (NER) tries to identify strings that represent names of e.g. persons, objects and events and annotate them in running text.

If a string is a name, it is often also interesting to know whom the string refers to in real life. If we find "Mannerheim", we may wish to know if it refers to the famous war hero and former President of Finland or to his sister and founder of the Mannerheim Children's Hospital.

For disambiguating references to named entities in real life, libraries use authority lists with ID numbers. An authority entry contains various details which can be used for disambiguating the reference in context. In addition, an authority list often contains attested spelling variants. Digitalization of documents using optical character recognition adds to the spurious spelling variants.

In this workshop we will get acquainted with NER for facilitating data mining and retrieval of documents in digital humanities using language technology.

Workshop

The presentations of the workshop deal with approaches to named entity recognition used in different Nordic countries as well as/or with how NER is applied / how to apply NER in history, literary research, political science etc.. Place: Unioninkatu 40, 3rd floor, room 15 (C306), Helsinki. Time: 9-10 June 2015.

Participants from outside Finland are advised to get in touch with their national CLARIN coordinators before registering.

Registration

Speakers, please register your details for the workshop using [this form](#) by Friday, May 15. Local participants, please register with [this form](#) by Friday, May 22.

Programme:

	Tuesday June 9
12.00 - 12.30	<i>Light refreshments</i>
12.30 - 12.50	Welcome and introduction
	Session on how NER is applied / how to apply NER in history, literary research, political science, etc.
12.50 - 13.10	Eeva Murtomaa (National Library of Finland): Authority Lists in Library Catalogues. Link to the slides
13.10 - 13.30	Eetu Mäkelä (University of Helsinki): Named Entities in Digital Humanities. Link to the slides
13.30 - 13.50	Dimitrios Kokkinakis (University of Gothenburg): Named entities in Swedish literature. Link to the slides
13.50 - 14.40	<i>Lunch</i>
	Session on approaches to named entity recognition used in different Nordic countries
14.50 - 15.10	Bjarte Johansen (University of Bergen): Named-Entity Chunking and Recognition for Norwegian. Link to the slides
15.10 - 15.30	Dimitrios Kokkinakis (University of Gothenburg): Annotation and evaluation of Swedish multiword named entities. Link to the slides

15.30 - 16.00	Miikka Silfverberg (University of Helsinki): Reverse Engineering a Rule-Based Finnish Named Entity Recognizer. Link to the slides
	<i>End of day</i>
19.00 -	<i>Dinner</i>
	Wednesday June 10
	Session on approaches to named entity recognition used in different Nordic countries
09.00 - 09.20	Dorte Haltrup Hansen (CST, Denmark): CST's Ruled based NER for Danish. Link to the slides
09.20 - 09.50	Marte Horndalen Howlid (The Norwegian Broadcasting Corporation, NRK): Migrating metadata - the need for NLP in NRK. Link to the slides
09.50 - 10.10	<i>Light refreshments</i>
10.10 - 10.30	Jyrki Niemi (University of Helsinki): Korp and NER. Link to the slides
10.30 - 10.50	Heikki Keskustalo (University of Tampere): Coping with OCR errors in NE search. Link to the slides
	Session on how NER is applied / how to apply NER in history, literary research, political science, etc.
10.50 - 11.10	Claus Povlsen (CST, Denmark): NER in real life!. Link to the slides
11.10 - 11.30	Dana Dannélls (University of Gothenburg): Geographic visualisation of place names in Swedish literary texts. Link to the slides
11.30 - 11.50	Discussion
11.50 - 12.30	<i>Lunch</i>
12.40 - 13.10	Demos – organised like a poster session, open to anyone who wants to demonstrate
13.10 - 14.00	Final session. Conclusions, actions etc. Krister Lindén: Summary of major outcomes and conclusions from the meeting. Next events. End of meeting

Accommodation

There are several hotels within walking distance, e.g.

- **Hotel Arthur** (***) , located very close to the workshop venue.
- **Hotel Cumulus Kaisaniemi** (****) , just a block away (might be somewhat noisier than Arthur).
- **Omenahotels** , a budget alternative, at walking distance but a bit further away.
- **Solo Sokos Hotel Torn**i (****) , for some luxury and spectacular views, at walking distance.

See more hotels in the center of Helsinki

Questions about the practical arrangements?

Please contact the organizers at [fin-clarin \(AT\) helsinki.fi](mailto:fin-clarin(AT)helsinki.fi) in case you have any questions.



Comments

Ask or comment!

In order to post a comment, you need to [register](#) and [log in](#).

Search the FIN-CLARIN wiki pages:

Search !

The Language Bank of Finland is brought to you by FIN-CLARIN. CSC is responsible for the technical maintenance in cooperation with the University of Helsinki.

© 2015 The Language Bank of Finland, FIN-CLARIN and CSC – IT Center for Science

[Log In](#) at [KitWiki](#)