

BiographyNet Meeting - Thursday, April 24th 2014, 14.00 - 15.15h.

Present: Milena, Piek, Niels.
On Skype: Serge.

Report: Niels

Data Hackathon 2 April (Niels)

Niels reported on the Data Hackathon event which he co-organiser with Victor de Boer, Chris Dijkshoorn and Albert Meroño Peñuela. The idea behind this data hackathon was to spend one day to create as many links as possible (within that time space) between the various (cultural heritage) data sets available within the VU Semantic Web group. The main reason was that if we advocate the advantages of Linked Data such as linking to other sources, we should at least do that between our own data sets and possibly write a paper on how easy/difficult this was and what kind of issues came up. With respect to BiographyNet, the main outcome was the creation of 5000+ links from people in the BiographyNet RDF data to people in the Rijksmuseum RDF data. This was done using automatic generation of skos:closeMatch links between people in BiographyNet and the Rijksmuseum Amsterdam collection data sets, using a SPARQL query in ClioPatria. The confidence for these links was determined as 50/50, as these links were made using string matching on names only. Other properties such as birth- and/or dead dates could be used to increase confidence, however within the time limitations of this event further improvements were not carried out. Nonetheless, this event really showed the great potential for linking the BiographyNet database to other data sources, which is one of the main outcomes of this event for BiographyNet. More info about the day and the results of other participants can be found at [\[1\]](#).

Workshop VU 4 April (report, Antske & Serge)

The workshop was considered to be a success. The detailed report can be found on [\[2\]](#). During the meeting Serge explained that most participants were positive about the workshop and stated to really have learned something about technologies we want to use such as RDF and the concepts of the Semantic Web. Some were a bit disappointed because they thought to learn more in terms of their actual data, i.e. find interesting new information for their research among their data. Though this was not communicated as one of the goals, it was also not explicitly stated that it was not. In that sense, communication could have been a bit more clear. Niels added that his main observation was that the participants (as most users of information systems) think about the Biografisch Portaal in a black-box fashion; in that sense it was hard for them to think of the actual data and the interface of the current BP as two separate entities. Furthermore, a lot of the feedback had to do with limitations of the current BP interface which to us were already sort of clear, but it was good to have a confirmation of the fact that a lot of the ideas we have for the BiographyNet demonstrator interface would have solved mentioned issues.

Presentation Semantic Web group

After giving a presentation about BiographyNet at the CLUE - Network Instituut day, Frank van Harmelen suggested to give a similar presentation to the semantic web group. Due to absence, the status of this topic is not completely clear. Niels will contact Victor to check if presentation during one of the SemWeb meetings was already discussed, and get some available slots in order to match with our agenda's.

CHAT (BiographyNet description; proposal Marjolein 't Hart)

This topic concerns the proposals for the Centre for Humanities And Technology, in which VU, UVA, IBM and eScience participate.

Piek mentioned that Antske would write a project description for BiographyNet. There are now over 50 submissions. There was an effort to cluster some of the submissions, however this effort has failed.

Running the current BP on our RDF corpus using ClioPatria? (Niels)

Niels mentions that a little more than a week ago, on an otherwise totally unrelated (semi-private) event, he met Jeltje Zijlstra from Huygens ING who (among other things such as

working on the '1001 Vrouwen' book) works on the Biografisch Portaal. During the talk that followed, Niels expressed to her that we are happy with the data from the BP and would like to find a way to include future updates and additions to the BP, as we currently only have our initial BP dump. However she expressed that there most likely are no further updates because, due to IT related difficulties, they are not able to work on the BP, which obviously is not a desirable situation.

Niels thinks that technically, it would not be difficult to let the current BP run on our RDF corpus using ClioPatria. This would be beneficial to Huygens ING as people would be able to continue their work on the BP, the technology behind the BP would be more 'future-proof', the current RDF data-model does not compromise concerns of data preservation and source-originality. It would be beneficial to the BiographyNet project as the RDF corpus will stay up to date and no further (manual) synchronisation needs to be done with future XML dumps of the BP (which have again to be checked for inconsistencies already solved in the then current RDF corpus). Also, it could be classified as 'low hanging fruit' for the project as such a solution would already contribute to the communities supporting BiographyNet without the full demonstrator being available. However, Niels can not say anything about political or organisational issues that may play a role. Piek mentions that Susan, Guus, he and people from Huygens ING would need to talk about this and were also planning to meet to discuss similar issues. The consensus is to move the topic to the agenda for the next meeting and try to have the relevant people present.

Papers-papers

The most recent publications related to BiographyNet are 'BiographyNet: Methodological Issues when NLP supports historical research' and 'NAF and GAF: Linking Linguistic Annotations'. Both will be presented in Iceland by Antske in the beginning of May. Serge mentioned that our journal article 'Biography and Digital Humanities: the history of the future?' is now under review at TSEG (Journal of Social and Economic History, [3]). Niels mentioned he is working on a Doctoral Consortium paper for ISWC 2014, concerning mainly his PhD but of course with ties to BiographyNet. The publication related to BiographyNet can be found at [4].

Absence Serge

As Serge's spouse is planned to deliver somewhere next week, Serge's absence thereafter was discussed. Serge plans to take up most of his leave from that moment on. As he is not 'out of the country' he will potentially be reachable for urgent matters and will try to follow developments, both using email.

Any Other Business (Rondvraag)

Milena announced that the current employer of their prospective candidate for coordinator eHumanities offered a contract extension, and that as such the candidate is no longer available. As there was no other suitable senior candidate available, a current colleague will partly and temporarily take over some of the coordination. Because taking over all of this work would not fit besides current obligations, Milena will continue the coordination of BiographyNet for eScience until further announcement.

Piek mentioned that he was contacted by Charles vd Heuvel of the 'Mapping Notes and Nodes in Networks' project [5][6] for possible collaboration. There is a lot of interesting data in the project, hence collaboration would be interesting at the data level as well as on other subjects such as the modelling we currently use for BiographyNet. However we should also not forget and focus on our own primary BiographyNet deliverables.

Serge had no further issues.

Niels asked for the ideas and status of the plans to create a roadmap for BiographyNet. All participants of the meeting considered this an important issue. Antske will be asked (upon return from holidays) about notes of these plans as discussed during the previous meeting. Current meeting participants think it is a good idea to use whatever is available now as starting point for a commonly shared document, e.g. a Google Doc, to further create and shape the roadmap, so we can all look at it and work on it.

Piek mentioned the upcoming Newsreader Hackathon in London [7] and invited Niels to join. Milena stated that perhaps this event would also be very interesting to eScience engineers and people working on other eScience projects. Piek thought this was a very good idea and is going to forward info on the event.

-/- End of meeting -/-

- [1] http://wiki.cs.vu.nl/biographynet/images/f/f6/Digitalhistorydatathon_02-04-2014.pdf
- [2] <https://wiki.cs.vu.nl/biographynet/images/6/6e/RapportageWorkshopBiographyNet.pdf>
- [3] <http://www.tseg.nl>
- [4] <http://www.biographynet.nl/publications/>
- [5] <http://www.huygens.knaw.nl/mapping-notes-and-nodes-in-networks/>
- [6] <http://acuh.uva.nl/home/components/news/news/content/folder/2014/03/new-project-mapping-notes-and-nodes-in-networks.html>
- [7] <http://www.eventbrite.com/e/kick-off-newsreader-and-hack-100000-world-cup-articles-tickets-2848605255?aff=eivtefrnd>