Digital Humanities Clinics – leading Dutch librarians into DH

Lotte Wilms*, Michiel Cock** and Ben Companjen***

*Koninklijke Bibliotheek, National Library of the Netherlands, Prins Willem-Alexanderhof 5, 2595 BE Den Haag, the Netherlands **Vrije Universiteit Library, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands ***Leiden University Libraries, Witte Singel 27, 2311 BG Leiden, the Netherlands

*lotte.wilms@kb.nl **m.p.cock@vu.nl ***b.a.companjen@library.leidenuniv.nl

Abstract. Academic libraries in the Netherlands have the ambition to increase the knowledge level of their librarians on digital humanities (DH). Three libraries therefore set up a series of full-day training events aimed specifically at library professionals of academic and research libraries in the Netherlands, named DH Clinics. The aim of these clinics is to provide basic methodological competencies and technical skills in DH, for a diverse group of library employees, consisting of both subject and technical librarians with basic technical skills. The content of these sessions should 1) enable them to provide services to researchers and students, 2) identify remaining gaps in knowledge or skills that they could address by self-directed learning and 3) (possibly) to automate their daily library work. This paper describes final schedule of the DH Clinics, how we organised the design process by involving the user community and the lessons we learned.

1. Introduction

Library professionals of academic libraries in the Netherlands have mostly worked with digital humanities (DH) researchers on an ad-hoc basis. Most organisations had not yet formed policy relating to digital humanities when the libraries came in contact with each other at a meeting in 2015. During this meeting, initiated by the KB, National Library of the Netherlands, the Vrije Universiteit Library and the Library Lab of Ghent University and inspired by the dh + lib community of the Association of College and Research Libraries Digital Humanities Special Interest Group[1], library professionals from the Netherlands and Belgium shared their activities related to DH and agreed to keep sharing experiences through a Google Group. However, due to changes in personnel, the Google Group was only initiated in 2016 and little had changed with regard to activities in the field of digital humanities within the academic library community.

Two members, Lotte Wilms, Digital Scholarship Advisor of the KB and Michiel Cock, at the time team manager Academic Support for Humanities and Social Sciences at the University Library of the Vrije Universiteit Amsterdam, were working independently on training activities on DH for their colleagues and decided to combine their efforts to organise sessions that would be open for all library colleagues from the GLAM community. Together, we proposed our idea for a national training programme to the board of UKB[2], the Dutch consortium of university libraries and the national library, for institutional support. Our proposal was

---

[1] dh + lib
[2] UKB

accepted and Ben Companjen, Digital Scholarship Librarian at Leiden University Libraries' Centre for Digital Scholarship was put forward by the director of his library to join us in our endeavour.

Together we started organising what we then named DH Clinics, the Dutch initiative to educate GLAM librarians on the digital humanities.

2. Inspiration

We were inspired by several other initiatives: Library Carpentry, the Programming Historian and the Digital Scholarship training programme at the British Library. Library Carpentry "introduces you to the fundamentals of computing and provides you with a platform for further self-directed learning"[3] and is aimed at librarians. The Programming Historian "offers novice-friendly, peer-reviewed tutorials that help humanists learn a wide range of digital tools, techniques, and workflows to facilitate their research."[4] And last, the training programme of the British Library "engages with those operating at the intersection of academic research, cultural heritage and technology to enable innovative use of [their] digital collections, and creates opportunities for library staff to develop skills necessary to support emerging areas of modern scholarship, particularly the Digital Humanities[5]".

Library Carpentry and the Programming Historian are great resources for tutorials on working with data using available software and on programming simple new applications for a specific task. The materials provided by both platforms have been collaboratively created and peer-reviewed. Library Carpentry, like its 'sibling-projects' Software Carpentry and Data Carpentry, provides not just lesson content, but lesson plans as well. The Digital Scholarship training programme run at the British Library showed that librarians not involved with digital tools or scholarship on a daily basis enjoyed the training and feel they benefit from it.

3. Goals

Our goals for the program are threefold:

1. to teach academic librarians about the basics of digital humanities
2. to strengthen the relationship between libraries and the academic community
3. to build a knowledge network of Dutch librarians working in DH

Our first goal is to provide librarians with basic methodological competencies and technical skills in DH. The content of these sessions should enable them to provide services to researchers and students, identify remaining gaps in knowledge or skills that they could address by self-directed learning and (perhaps) to automate their daily library work by using the techniques the clinics introduce them to. We are not setting out to turn them into programmers or data crunchers, but want to boost their knowledge level to where they feel comfortable providing information about DH projects, follow the literature and research,

---

[3] Library Carpentry
[4] The Programming Historian
[5] McGregor et al. (2016)

follow online tutorials and hopefully take up the challenge of continuing this professional development by engaging with the DH community.

Second, we wish to provide the library professionals access to a network of digital humanities researchers within and across universities and research institutes, by asking those researchers to teach the sessions in the clinics. We also hope that learning about the computational possibilities in humanities research directly from the researchers themselves, who often talk about their work with great enthusiasm, will spark the librarian's interest in these matters.

Last, we hope the program will provide an opportunity for librarians to connect and build a knowledge network across institutional boundaries, both offline during the sessions and online (e.g. through the aforementioned DH+LIB NL/VL Google Group). If the program is successful, we will maintain the community feeling by organising follow-ups, e.g. in the form of a DH Project Day, in which we collaboratively work on a library related DH project, such as cleaning and linking catalogue data.

4. Selection of topics

Digital humanities is a term that is difficult, if not impossible, to define using a limited number of research topics. However, a certain selection of topics was needed to decide the scope of the clinics. Not wanting to miss a relevant method or application, we started our selection process with a brainstorm with several librarians working in DH. This produced a list of concepts, methods and software applications. We then compared this list to other educational programmes in DH, most notably The Programming Historian and Johanna Drucker's Introduction to Digital Humanities[6], to make sure we had not missed any important items, which as it happened we did not.

Next to this, we did a literature review on digital humanities in libraries. What does it entail to be a DH librarian and what skills are required, which are currently lacking?[7] Two surveys done in the United States provide a good overview of skills needed for librarians who work in the digital humanities. The Association of Research Libraries published the report 'Supporting Digital Scholarship', written by Rikk Mulligan who surveyed 73 ARL member libraries. Molly Dahl Poremski undertook a similar endeavour and surveyed 117 librarians who agreed with the statement "I am a librarian that works with digital humanities[8]" in her study. The two reports found similar results with regards to skills on DH for librarians. As Poremski states, "it is the technologically advanced skills (digital mapping, text encoding, and computer programming) that are most often lacking in the current digital librarians' skill set"[9]. This was also found by Mulligan who saw that "[t]he greatest gaps remain in visualization (65%), computational text analysis and support (64%), statistical analysis support (60%), and in developing software (54%)."[10] These US based surveys are supported

---

[6] Drucker et al (2014)

[7] For all collected literature, see
https://www.zotero.org/groups/969489/liber_digital_humanities_working_group/items/collectionKey/WXP2WW4G

[8] Poremski (2017) 4

[9] Poremski (2017) 10

[10] Mulligan (2016) 8

by the recently published report by Christina Kamposiori of Research Libraries UK (RLUK)[11], who undertook a similar survey amongst 27 libraries, of which 26 were RLUK member organisation and found that those surveyed looked for skills in DH librarians such as programming (77.8%), a knowledge base (77.8%), negotiation and communication skills (66,7%), leadership skills (55,6%) and research skills (59,3%)[12].

Next to this, we gathered 15 job descriptions to get an overview of the qualifications that libraries ask of their DH librarians (see Figure 1).
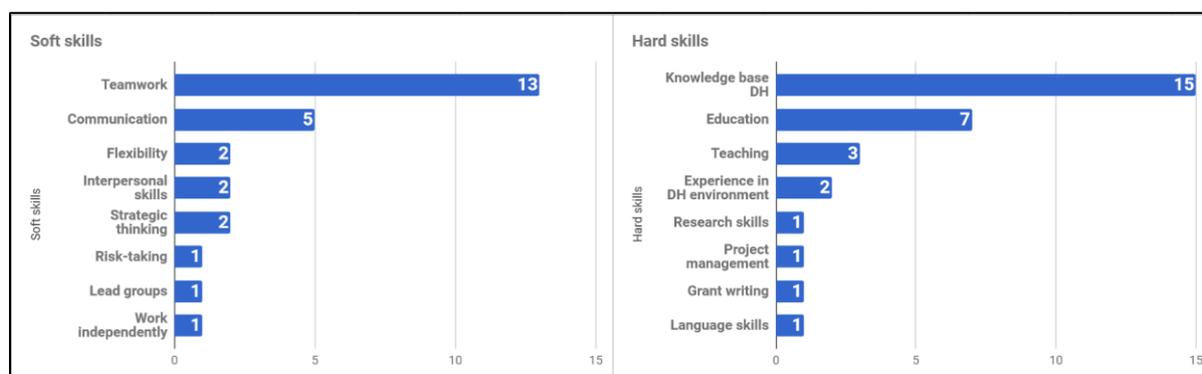

Figure 1 – Soft skills versus hard skills in job descriptions (n=15)

From this we learned that soft skills such as collaboration, communication and strategic thinking are asked quite often from a DH librarian, but that a general knowledge of digital research methods and tools returns in all descriptions, in one form or another, as is confirmed in the RLUK survey above. This overview is exactly what we intend to provide the Dutch librarians with the DH Clinics.

We believed it was important to not only address the topics we came across in our work as librarians, but to also see what researchers needed and expected from their libraries. We therefore took this discussion to the THATcamp Amsterdam on 27 October 2016 where we proposed a session on 'DH and Libraries'[13]. The accepted session gave us the opportunity to discuss our ideas and list of topics with researchers from the DH community. This provided us with very useful feedback, both on the role of a DH librarian in general and on our ideas for the programme. The most important advice was to focus on teaching the library professionals about concepts and methods, not specific tools. One specific topic that participants felt librarians did not understand well enough was the influence of data completeness and correctness on DH research, which has been a reason to include data criticism in the programme.

5. Organisation of topics in sessions

Our list of concepts, methods and software applications was still a very large one, without any self-evident structure that would help us divide the list into training sessions. We therefore mapped the items on our list onto the TaDiRAH[14] research taxonomy (Taxonomy

---

[11] Kamposiori (2017)
[12] Kamposiori (2017) 16
[13] Cock and Claeyssens (2016)
[14] TaDiRAH - Taxonomy of Digital Research Activities in the Humanities

of Digital Research Activities in the Humanities). To do this, we also used a subdivision of our topics related to text, structured data and audio visual material.[15] Some general topics we assigned to multiple phases in the taxonomy, e.g. Natural Language Processing (which we assigned to both Enrichment/Annotating and Analysis/Structural Analysis) and XML (which we assigned to both Capture/Transcription and Enrichment/Editing). These borderline cases were few and most of our items could be assigned to one phase.

Looking at this distribution over topics of TaDiRAH and the current practice of libraries in the Netherlands, we decided to select four phases which we felt were underrepresented in academic libraries: capture, creation, enrichment and analysis. Due to the diversity of material we wanted to include and the different approaches needed, we chose to subdivide analysis into text and non-text.

6. Library community feedback

To ensure the topics we selected from the list are also what the participants expect from the clinics, we invited them to offer feedback on our programme. We put together a survey using Google Forms, in which we asked respondents to indicate their preferences by ranking three possible workshop topics workshop for each day.

Each question had a short introduction about the topic of the day, the proposed workshops and links to more information. The respondents could then select their first, second and third choice of workshop (Figure 2 and Figure 3). We also asked them if we had missed a crucial topic in our programme. Next to this, we asked more general questions about the organisation, such as which day they would prefer the lectures to be on and whether they would like to follow the whole programme or just a (number of) day(s) (see Figure 4).

---

[15] Based on the division of the Dutch research infrastructure consortium CLARIAH (the Common Lab Research Infrastructure for the Arts and Humanities, https://www.clariah.nl) into three work packages

Dag 1: Digitalisering en databases

Deze dag draait om het digitaliseringsproces en het maken en gebruiken van databases.

SQL is een taal om databases mee te bevragen. Meer informatie: https://en.wikipedia.org/wiki/SQL

SPARQL is een variant op SQL en wordt gebruikt voor om RDF-databases mee te bevragen. Meer informatie: https://en.wikipedia.org/wiki/SPARQL en https://en.wikipedia.org/wiki/Resource_Description_Framework

De introductie over data behandelt het jargon dat wordt gebruikt bij het spreken over data, code en software. Ook kijken we naar welke methoden je het beste kunt hanteren als je werkt met data. Meer informatie: https://data-lessons.github.io/library-data-intro/

	1e keus	2e keus	3e keus
SQL	○	○	○
SPARQL	○	○	○
Data introductie	○	○	○

Figure 2 – Example of survey question
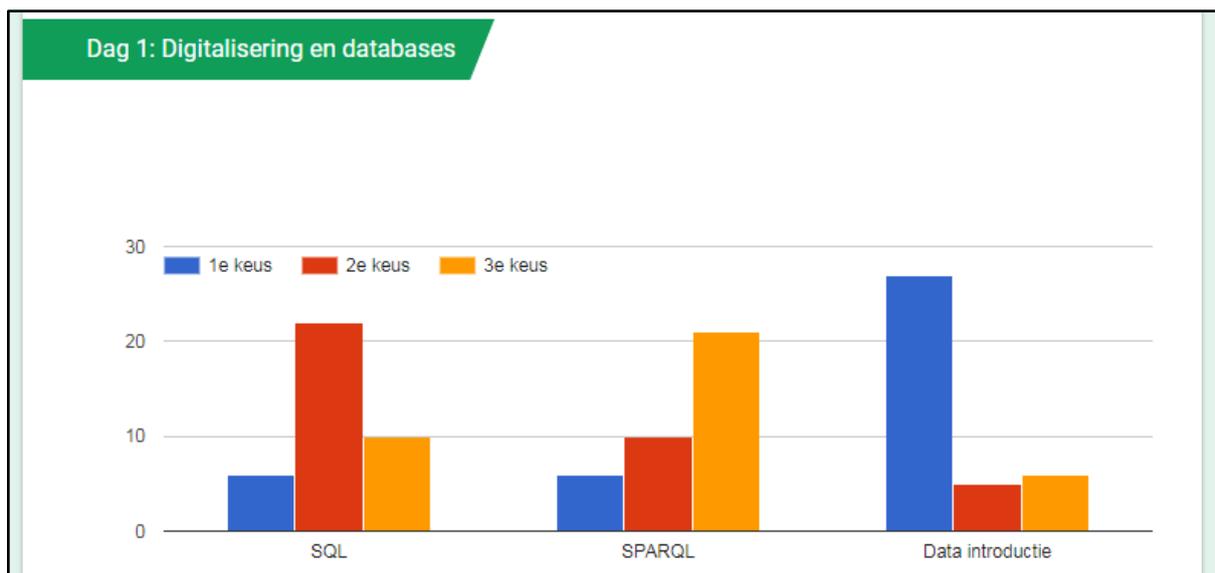
Dag 1: Digitalisering en databases

Figure 3 - Example of collected feedback. Blue is first choice, red is second choice and light orange is third choice.
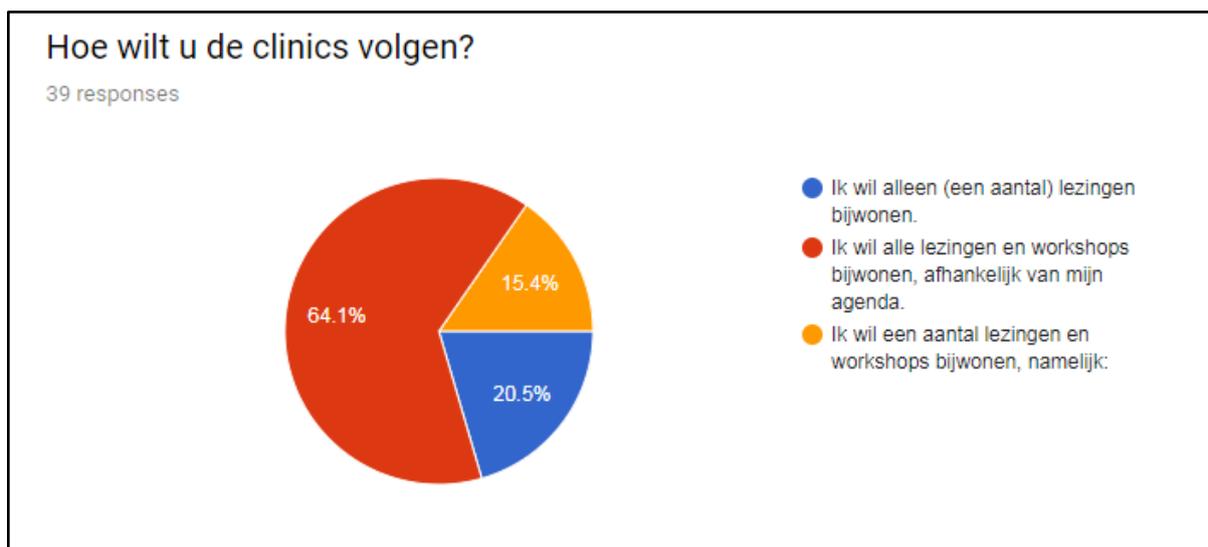
Figure 4 - Response to question: 'How do you want to follow the clinics?'. Blue is 'I only want to join you for (a few) lectures', dark orange is 'I want to do all lectures and workshops, depending on my agenda' and light orange is 'I want to do some lectures and workshops'.

The survey was sent out to members of the Google Group DH+Lib NL/VL, colleagues in the GLAM community and posted as a news item on the KB Lab website[16], which was shared in its newsletter[17] and on Twitter[18].

Within the two weeks that the survey was open (25 April - 10 May 2017) we received 40 responses, from 14 institutions (see figure 5 below). The timing of the survey was not optimal, as it was the national May holiday during these two weeks and we received several messages from people who would have liked to provided us with feedback, but were away from the office.

---

[16] Wilms (2017)
[17] KB Lab newsletter (2017)
[18] KBNLResearch (2017)

Count of Institution

Leiden 17.5% — 7
VU 20.0% — 8
Gent 2.5%
EUR 2.5%
Independent 2.5% — 1
KB 22.5% — 9
UU 10.0% — 4
CERL 2.5% — 1
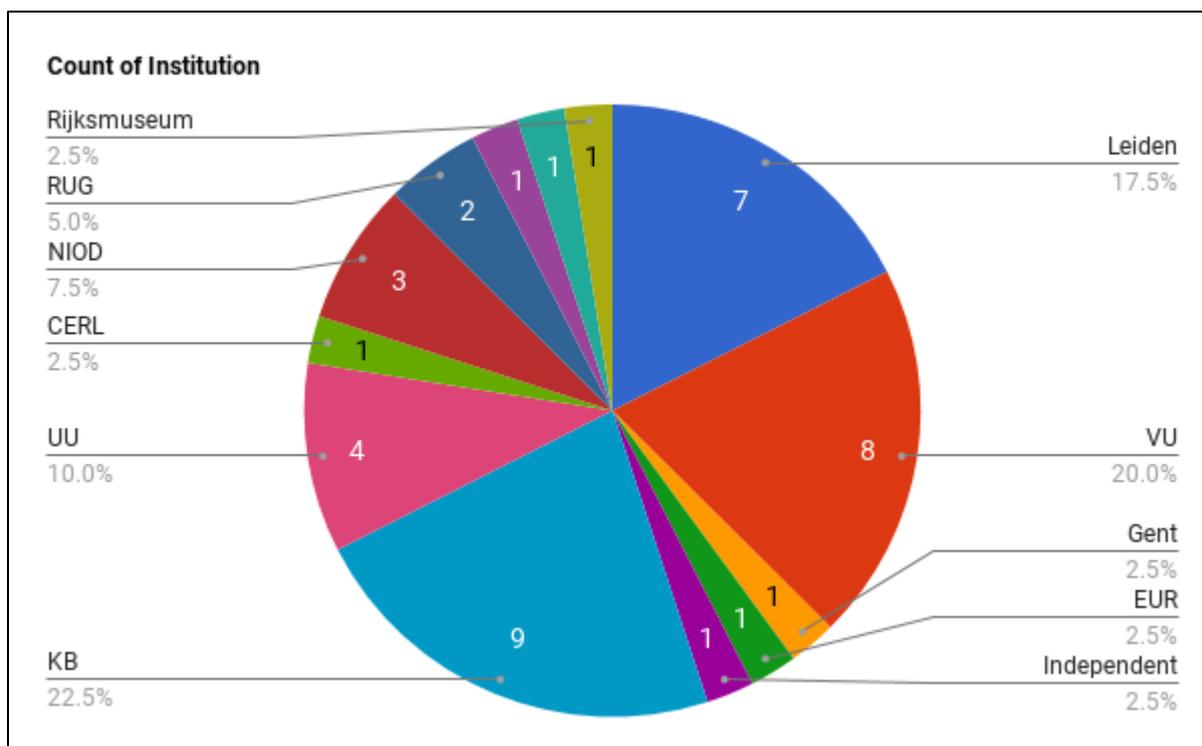NIOD 7.5% — 3
RUG 5.0% — 2
Rijksmuseum 2.5% — 1

Figure 5 - Number of respondents of the survey and corresponding institutions

The feedback provided in the multiple choice questions gave us a mostly clear overview of what respondents are interested in. The open questions led us to adapting some of the workshops we suggested. For example, one respondent added that they would like to know more about how to model data for a database, which was not something we had provided as an option.

Other responses to our open question on what we had missed in the survey will be addressed during the whole programme, such as examples based on current library practice, information about often used software and background information on digital humanities.

7. Setup

Once the feedback cycle was complete, we could finalise the programme and start the practical organisation of the events. The clinics are scheduled as full-day events with three weeks between each event, from 5 September 2017 until 7 November 2017, organised in central locations in the Netherlands. Each morning session consists of two lectures and the afternoon sessions are practical, hands-on workshops.

To ensure we will reach the widest audience possible, the afternoon sessions are not mandatory. Librarians who only want to know about DH, but have no interest in or time for working with the tools and methods can simply join the lectures, while attendees who wish to be able to work with applications can participate in the workshops. We do require the workshop participants to also attend the lectures in the morning.

The programme is organised in such a way that each clinic is a separate entity, so participants do not have to follow the whole five-day programme. They can pick and choose what they find relevant for their work.

From the survey we conducted, we expect about 30-40 participants for the lectures and 15-20 participants for the workshops, which is also the maximum number of people we can accommodate in a workshop setting.

8. Lecturers

To find lecturers for the morning sessions, we wrote profiles for each lecturer, such as;

> Is (or was) responsible for a large digitisation project and can share knowledge on processes, copyright, techniques and data formats. Added bonus would be if the person also has experience with digitisation of audio, video and objects.

With these profiles we contacted the CLARIAH consortium for tips to find lecturers that fit our profiles. With their suggestions and our combined knowledge of the DH field we contacted several people from not only universities, but also research institutes throughout the Netherlands and Belgium. We also paid attention to having a good gender balance and that the people we contacted already gave similar lectures before to ensure the time they needed to prepare would be limited.

The programme of the workshops was finalised after the aforementioned survey was closed. We believe it to be very important to reuse what is available and looked for teachers who would be able to use already existing material, such as the Library Carpentry workshops. Here, we took the same approach as with the lectures and contacted people who already gave similar workshops or were familiar with the material available online.

9. Schedule

With the final schedule we hope to address several main activities of and related to the digital humanities, which are relevant for library professionals.

	Lecture 1	**Lecture 2**	**Workshop**
Day 1 - Capture	Digitisation	Databases	SQL for Librarians
Day 2 - Creation	Computational thinking	Code, tool and source criticism	Introduction to Python
Day 3 - Enrichment	Natural language processing and Named entity recognition	Linked Data	Cleaning data
Day 4 – Analysis text	Text analysis with historical sources	Text analysis with born-digital sources	Network analysis

Day 5 – Analysis non-text	Analysis of audio visual sources	GIS	Computer vision

We intend to incorporate feedback we receive from the participants during the first clinics in further clinics.

10.   Practicalities

To keep an overview of all activities and training materials, we set up a Github Pages website at dhclinics.github.io. From this website we link to the registration forms on Eventbrite[19]. Due to the limited spaces available in the workshops, we ask that no more than two people from one organisation join the workshop.

We do not have our own budget, but work with day sponsors. Libraries and supporting organisations are asked to pledge a small sum (approximately 500 euro) to cover costs for lunch and travel costs for lecturers and workshop instructors. In return they receive 5 minutes of speaking time at the start of their sponsored day and their logo on all online and offline communication of that day.

11.   Lessons learned

Even though the clinics have yet to take place, we have already learned much from the design process. As often, a major lesson is that it is very important to decide your scope. Try to limit it as much as you can, to make sure you can focus on what is important for your cause. We did this for example when deciding which categories to select for a theme. If there was already a good training programme in place within the Netherlands, we chose to not duplicate that.

We also learned that it was quite easy to involve the community that we were working for. The survey was put together quite quickly and respondents were not difficult to find. However, what was more difficult to find was time. It is important to ensure you have enough free time to work on an effort as big as this, because otherwise things move too slowly. We therefore planned weekly Skype calls with each other and used a Trello[20] board to keep track of what was happening and make sure that we spent the necessary time on the actions we discussed.

By getting together virtually on a weekly basis, we got to keep the speed of our efforts, but it is important to not be too quick. We wanted to organise a first clinic before the summer, but it soon became apparent that we could not get the speakers we needed for this date. We therefore decided to plan all clinics in the fall and winter. Do not be afraid to postpone if that improves your sessions.

---

[19] Eventbrite
[20] Trello

12. Future plans

Registrations for the clinics open in the summer and we will be continuously talking to our participants to improve the following clinics and our communication.

When the DH Clinics are successful we plan to organise a DH Project day (or two) where all participants join forces to design a small DH project together, ideally with a library base, such as catalogue data. We will formulate a research question, acquire source material, clean the data, analyse the data, answer questions and share the results. With this, we hope the participants will be able to put what they learned into practice and to not only provide support in the endeavours of the researchers they work with in their libraries, but also possibly improve their own workflow by applying the method they learn in this project, such as cleaning data, to their daily work.

References

CLARIAH - Infrastructure for the Humanities. Available at https://www.clariah.nl, accessed 08.08.2017.

Cock, Michiel; Claeyssens, Steven (2016): THATCamp Amsterdam: Proposal for a Talk Session: DH and Libraries. Available at http://amsterdam2016.thatcamp.org/2016/10/19/talk-session-dh-and-libraries/.

dh + lib. Available at http://acrl.ala.org/dh/, accessed 08.08.2017.

Drucker, Johanna; Kim, David; Salehian, Iman; Bushong, Anthony (2014): Introduction to Digital Humanities: Concepts, Methods, and Tutorial for Students and Instructors.

Eventbrite. Available at eventbrite.com, accessed 08.08.2017.

Gitter. Available at https://gitter.im/, accessed 08.08.2017.

Kamposiori, Christina (2017): The role of Research Libraries in the creation, archiving, curation, and preservation of tools for the Digital Humanities. Research Libraries UK. Available at http://www.rluk.ac.uk/wp-content/uploads/2017/07/Digital-Humanities-report-Jul-17.pdf. Publication date July 2017.

KB Lab newsletter (April 2017). Available at http://lab.kb.nl/newsletter/2017-04/, accessed 08.08.2017.

KBNLResearch (2017): Tweet. Available at https://twitter.com/KBNLresearch/status/857192078899544064, accessed 08.08.2017.

Library Carpentry: Software skills for library professionals. Available at https://librarycarpentry.github.io/, accessed 08.08.2017.

McGregor, N.; Ridge, M.; Wisdom, S.; Alencar-Brayner, A. (2016): The Digital Scholarship Training Programme at British Library: Concluding Report & Future Developments. In *Digital Humanities 2016: Conference Abstracts,* pp. 623–625. Kraków: Jagiellonian University & Pedagogical University. Available at http://dh2016.adho.org/abstracts/178, accessed 08.08.2017

Mulligan, Rikk (2016): SPEC Kit 350: Supporting Digital Scholarship. Association of Research Libraries. Available at http://publications.arl.org/Supporting-Digital-Scholarship-SPEC-Kit-350. Publication date May 2016.

Poremski, Molly Dahl (2017): Evaluating the landscape of digital humanities librarianship. In: *College & Undergraduate Libraries,* 3 (4), 1–15. doi: 10.1080/10691316.2017.1325721.

TaDiRAH - Taxonomy of Digital Research Activities in the Humanities. Available at http://tadirah.dariah.eu/vocab/index.php, accessed 08.08.2017.

The Programming Historian. Available at https://programminghistorian.org/, accessed 08.08.2017.

Trello. Available at https://trello.com/, accessed 08.08.2017.

UKB. Available at https://www.ukb.nl/english, accessed 08.08.2017.

Wilms, Lotte (2017): Digital Humanities Clinics voor bibliotheekmedewerkers: input gevraagd! Available at http://lab.kb.nl/news/digital-humanities-clinics-voor-bibliotheekmedewerkers-input-gevraagd, accessed 08.08.2017.