Preface by the Guest Editors

Ingrid Mayeur[1] and Claartje Rasterhoff[1]

[1]University of Liège
[2]University of Amsterdam

The sixth *DH Benelux Conference* was held on 11 — 13 September 2019 at the University of Liège (ULɪÈɢᴇ), Belgium. The event was organised under the auspices of the CIPL (computer centre of the Faculty of Philosophy and Letters,[1] directed by Dr. Björn-Olav Dozo) and the LASLA (Laboratory of Statistical Analysis of Ancient Languages,[2] directed by Prof. Dr. Dominique Longrée). During those three days, the conference brought together over a hundred participants around the theme "Digital Humanities in Society". Starting in 2014, the annual symposium *DH Benelux* aims to stimulate the collaboration between Digital Humanities researchers in Belgium, The Netherlands, and Luxembourg — although it remains open to everyone, including researchers from outside the Benelux[3]. The conference therefore presents an opportunity for the community of *digital humanists* to meet and exchange around intellectual (or even material) nourishment, by introducing their ongoing projects, discussing their results, and testing their tools. Building on a long tradition of dialogues between the humanities and the computer sciences, the University of Liège proudly hosted the 2019 edition. The two research centers involved in the organisation, the CIPL and the LASLA, have indeed long been concerned with the development of Digital Humanities in Belgium. The first was created in 1983, aiming to promote and to coordinate the use of computer science within the Faculty of Philosophy and Letters. The second was founded earlier, in November 1961, and was the first research centre to have studied the classical languages — Greek and Latin — using automatic information processing technologies. In doing so, the LASLA has collected in computer files numerous ancient Latin works, from Plautus to Ausone, as well as texts from classical Greek literature.

The background of the hosting institutes reflects the way in which, initially, Digital Humanities took off by putting computer technology at the service of research in the humanities —- also described as *Humanities Computing*, and illustrated by, for example, McCarty's eponymous book (McCarty, 2005). Digital technologies have since then constantly evolved, making it possible for humanities scholars to take into account new objects of study, to scale up data collection and to present results in new ways. Moreover, under the influence of web technologies and the networking of texts and data — or, more broadly speaking, of the widespread changeover of our societies to the *digital* — research in the humanities has acquired unprecedented possibilities of

---

[1]   `https://www.cipl.uliege.be/cms/c_4535714/fr/cipl`
[2]   `http://web.philo.ulg.ac.be/lasla/`. LASLA is part of the research unit (UR) *MOndes Anciens.*
[3]   See also the website of the event: `http://2019.dhbenelux.org/`

dissemination and interaction. Digital Humanities research now goes well beyond the use of computer tools in the humanities, by addressing the various ways in which the humanities are impacted by digitization and datafication, and the role they might be called upon to play in an increasingly digital society.

Although their worth is still too often contested, the humanities are involved in the production of heuristic and critical knowledge that enables actions in the social world as well as the very possibility of a democratic debate (Nussbaum, 2010, Small, 2013). Today, this social world, as well as the man-made artefacts that humanities scholars study, is increasingly digital and datafied (Doueihi, 2008). The changing practices of humanistic research under the impact of digital media, as well as the humanities' ability to question the materiality of the underlying digital infrastructures, challenge us to consider the socio-political make-up of Digital Humanities (Bonde Thylstrup, 2019, Mounier, 2018). The 2019 edition of the *DH Benelux* conference was therefore especially interested in research that addresses Digital Humanities in relation to broader societal transformations: whether these involve new forms of knowledge production and consumption such as citizen science and participatory research methods, or relate to processes of digitisation and datafication in society, including ethical and political issues. In that respect, the symposium aimed to open up the debate on how Digital Humanities should position itself in relation to the various institutional policies that fund or request research that engages with big data, artificial intelligence and data visualizations, and that encourage collaborations with both private and public partners.

Keynotes Lectures

The keynote lectures by Tim Hitchcock (University of Sussex) and Helle Strandgaard Jensen (Aarhus University) addressed the theme of the conference head-on. Both lectures addressed what digital processes in knowledge production and consumption mean for present-day humanities scholarship, and they both confronted us with what it means to be a responsible researcher. Strandgaard Jensen did so by bringing cultural theorist and political activist Stuart Hall (1932-2014) into the conversation, and Tim Hitchcock by invoking Sarah Durrant — a 61 year old widow who, in 1871, was charged with stealing two bank notes. Together, Hitchcock and Strandgaard Jensen revisited the notions of the library and the archive, respectively, employing historical and cultural analyses to create awareness of the political economies and technologies that shape our research on all levels. Their message was clear: when it comes to understanding the knowledge ecosystem in which we work and to which we contribute we need to do more and we need to do better.

Money, Morals and Representation. The day Stuart Hall joined my Archives 101 class

Strandgaard Jensen's lecture on digital archiving literacy reflected her efforts to raise the awareness of historians and other researchers with regard to the way in which digital processes invariably impact their work and their disciplines, and to help archival institutions understand the role they play in this process (Strandgaard Jensen, 2020). Indeed, when collections are being digitized they go through a process of remediation and become part of a new cyberinfrastructure – an infrastructure that too many researchers are still too unfamiliar with. When we (re)use the data from digital archives, she argued, we should do it wisely and knowingly – and that includes understanding the political economy and technical designs of digital archives. As Strandgaard Jensen

suggested, Stuart Hall's model can help with this, as it demonstrates how meaning is encoded into the cultural products we consume. As such, we can understand digital archives as digital objects that are encoded by librarians and researchers, funded by stakeholders, made accessible by policy makers, developed by web developers and software engineers, and even the technical capabilities and limitations of the medium they are developed in. But archival institutions also to a large extent conceptualize their archives with a specific user in mind — and in the case of digital collections, those users are often not researchers.

In applying Hall's model to the research of digital archives, Strandgaard Jensen theoretically and empirically researches the archive as a medium that gets remediated when its holdings are digitized. How then has the digital transformation of archival holdings and finding aids affected possibilities for data reuse? How can documentation help researchers avoid data misuse? Based on interviews, analyses of policy papers and the front ends of digital archives, and a multidisciplinary literature review, Strandgaard Jensen encouraged us to improve our collaborative practices between humanities researchers and archival institutions, and to invest in teaching digital (archival) literacy. Her lecture also demonstrated the added value of using the notion of the archive to understand and engage with digital infrastructures, and the need for more empirical research that addresses the construction and knowledge organisation of digital archives, and its impact on methodologies.

Visualising the Infinite Archive

In his lecture, Tim Hitchcock posited that our research methodologies have not kept pace with changing technologies, and that as a result, we now struggle to find trends and meaning in the masses of available data. He argued that there is a fundamental problem with the way in which we represent historical data, or more broadly humanities data, on our screens: with the way in which we search for data, and how we interrogate our search results. Hitchcock argued that the prototypical "lonely search box in the middle of the screen" of most of today's search engines symbolizes a tendency to hide information and strip data of its context — whereas it is exactly this dialogue between data and its sources that is key to effective scholarship. The first step in re-imaging humanities research, Hitchcock proposed, is to go back to the old idea of the library, to rethink our relationship with that "machine for knowing", and to acknowledge the power technologies (both old and new) hold in shaping our research. Here a "macroscope" approach (Börner, 2011), can help us re-imagine search, discovery and research, by providing a new form of "radical contextualisation".

The "macroscope", Hitchcock explained, allows you to see an object at all scales at once — from the most distant to the most granular. It thereby attempts to reconfigure the tools to match humanist methods and, at the same time, to reconfigure our representation of the library as an institution that helps us understand the knowledge systems within which we are working. In his presentation of some of the strategies he developed in collaboration with his colleague Ben Jackson, Hitchcock demonstrated how tools for textual and data analysis can be combined to re-invent a visible and visual context for data. In their demo, they positioned the Old Bailey Online dataset, which encompasses accounts of some 197,745 trials held at the Old Bailey in London between 1674 and 1913 in relation to a set of library and archival catalogues, with the purpose to "allow a new 'open eyed' way of working with data of all sorts — to allow macro-patterns and clusters to be identified; while single words and phrases can be

fully contextualised".[4] The value of this approach, then, lies not only in the possibility to combine close and distant reading, but also in using these technologies to expose the limits of our collections, as well as the structures of authority they reflect — and, by extension, the limits of our knowing.

Journal articles

The four articles selected for this issue are based on papers that were presented during the conference. They are of interest with regard to the conference theme "Digital Humanities in Society", either by providing through digital methods a better knowledge of the past in order to understand current social/cultural events, or by investigating the digital circulation of research objects specific to the Humanities. For the most part, these papers are the result of a collaborative work. An opportunity to demonstrate once again — if this is still necessary — that the Digital Humanities are a lively field that values the collaborative component of research work.

The contribution that opens this journal issue, "The Datafication of Early Modern Ordinances: Text Recognition, Segmentation, and Categorisation", directly echoes the issue raised by the keynote speakers of the digital valorisation of heritage texts. C. Annemieke Romein (Ghent University/University Rotterdam/KB National Library of the Netherlands), Sara Veldhoen (KB National Library of the Netherlands) and Michel de Gruijter (KB National Library of the Netherlands) report on the challenges they encountered in the *datafication* of a corpus of early modern printed normative texts (*i.e.* public ordinances or *placards*) under the project *Entangled Histories*. It addresses the need for software-based solutions for recognizing the complex Dutch Gothic print, the segmentation of texts compiled in books of ordinance, the creation of relevant categories of texts, and the automation of categorization. Even if this *datafication* serves to improve knowledge of the rules of Federation-State, such a feedback can be read as a sharing of good practices that could be applied to the treatment of similar collections.

Such *datafication* of old texts helps their automated processing and can result in a reevaluation of previously accepted ideas about these corpora. Theories and findings from other disciplines, then, can help scholars make sense of patterns in larger text corpora, as demonstrated in the contribution by Gianluca Valenti (ULiège): "A Corpus-Based Approach to Michelangelo's Epistolary Language" . In his essay, Valenti mobilizes quantitative methods such as correspondence analysis and correspondence regression on Michelangelo's entire epistolary corpus — about 500 handwritten letters. He investigates the traces of a language smoothing over time by using the theoretical frameworks of sociolinguistics and the abundant scientific knowledge of the Florentine dialect. The author shows that, although it is commonly asserted that Michelangelo's epistolary language would be close to the common contemporary language of 16th-century Florence, his letters display a tension between this language and that of the 14th-century Old Florentine tradition. And that from 1530 onwards — the time when Michelangelo reached the status of a public figure — forms from this Old Florentine language became increasingly prevalent.

The contribution of Chris Tanasescu (UCLouvain, Belgium), Diana Inkpen, Vaibhav Kesarwani and Prasadith Kirinde Gamaarachchige (all three from University of Ottawa) entitled "A-poetic Technology. #GraphPoem and the Social Function of Computational Performance" also exploits the opportunities of computational processing of literary corpora with a focus on its social and technical aspects. The #GraphPoem project

---

4   https://oldbaileyvoices.org/

relies on the hypothesis of a *performative networked sociality of poetry in digital culture* depending on both humans, poems and machines. The project intends to highlight the way in which poetic texts shape their environment and create the conditions for their reception as they are disseminated within digital media. Starting from such an assumption requires us *to go beyond the poetry*, and to investigate how *medial* and *computational* features actively forge the text as a poem in this digital context. The essay's scientific approach is based on a theoretical framework that integrates both the philosophy of Simondon's technique, and the reappropriation of von Uexkull's *Umwelt* concept by J. A. Schwarz. It leads to an algorithmic treatment of a corpus of digitized poems which intends to uncover the network of relationships in which they are intertwined. It also includes a participatory perspective involving the public in interactive digital performances of computational poetry in order to underline the social dimension of the writing-reading process of poetry through digital spaces.

Using network analysis and data processing tools responsibly means integrating concern for transparency, reliability and reproducibility of research results. The article of Julie M. Birkholz (Ghent University) and Albert Meroño-Peñuela (Vrije Universiteit Amsterdam) addresses this issue through the example of knowledge graphs using the Resource Description Framework language (RDF). These graphs are very popular among digital scholars since RDF provides structured/linked data on cultural objects that are readable by both humans and machines. It thus logically paves the way for network analyses. However, the authors point out the complexities encountered in such an approach — especially the risk of *black boxed tools* — and in making it explicit and reproducible. They therefore introduce a *proof of concept* relying on a concrete tool — a publicly accessible Jupyter Notebook that combines popular libraries in RDF data management and network analysis — the relevance of which they illustrate through two concrete case studies.

Acknowledgements

We end these lines at a time when preventive measures to contain the spread of covid-19 resulted in the cancellation of the *DH Benelux* 2020 live event in Leiden, that was then promptly replaced by a slimmed down online version of the conference. With a thought for all those who have seen their lives turned upside down by the pandemic in one way or another, we hope that the 2021 edition will help us to renew our tradition of scientific exchange in the Digital Humanities community in the Benelux in a fruitful way — a dialogue that we look forward to continuing in person once more.

References

Bonde Thylstrup, N.

2019. *The Politics of Mass Digitization*. Cambridge, MA: The MIT Press.

Börner, K.

2011. Plug-and-play macroscopes. *Communications of the ACM*, 54(3):60–69.

Doueihi, M.

2008. *La grande conversion numérique. [suivi de] Rêveries d'un promeneur numérique*, Points. Paris: Seuil. CR de Pierre Mounier: http://www.homo-numericus.net/article282.html Date de première édition: 2008.

McCarty, W.

2005. *Humanities Computing*. Basingstoke: Palgrave Macmillan. CR: http://www.digitalhumanities.org/dhq/vol/1/1/000001/000001.html.

Mounier, P.

2018. *Les humanités numériques : Une histoire critique*, Interventions. Paris: Éditions de la Maison des sciences de l'homme.

Nussbaum, M. C.

2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton: Princeton University Press.

Small, H. H.

2013. *The Value of the Humanities*. Oxford: Oxford University Press. CR: http://www.wsj.com/articles/SB10001424052702303650204579374740405172228 Première édition: 2013.

Strandgaard Jensen, H.

2020. Digital Archival Literacy for (all) Historians. *Media History*, 0(0):1–15. Publisher: Routledge _eprint: https://doi.org/10.1080/13688804.2020.1779047.