

What does that little black square store? The contents of Herman de Coninck's floppy disks in the Letterenhuis

Lamyk Bekius, University of Antwerp

lamyk.bekius@uantwerpen.be | <https://orcid.org/0000-0002-4533-2880>

Jordan Thijs, University of Antwerp

jordan.thijs@student.uantwerpen.be | <https://orcid.org/0009-0001-8504-5560>

Short paper

Computers have been a widespread writing technology since the popularisation of the word processor in the early 1980s (Kirschenbaum et al. 2009), but digital materiality is only slowly entering (literary) archival institutions because born-digital archives often remain in the private ownership of the author (Reside 2014). Still, when born-digital archives are part of the collection, they are often “uncatalogued, unfindable and unusable” (Jaillant 2022, 418). Research on born-digital archives is therefore scarce (Jaillant 2022; Ries & Palkó 2019; Reside 2014). The aim of the [CLARIAH-VL](#) Open Humanities Service Infrastructure is to advance digitally-enabled research in Humanities and the Arts by, among other things, providing data-level access to digitized and born-digital resources. This short presentation presents the ongoing efforts of one CLARIAH-VL's research scenarios: ‘Unlocking born-digital literary heritage: the case of the Herman de Coninck floppy disks’. This specific research scenario uses the CLARIAH-VL infrastructure to make born-digital literary archival material in Flanders findable and more accessible, starting with Herman de Coninck's floppy disks in the collection of the Letterenhuis (Antwerp).

Aim of the research scenario

The floppy disks of the prominent Belgian poet, essayist, journalist, and publisher Herman de Coninck (1944-1997) hold a special place in the Letterenhuis collection, as they were the first born-digital archival material acquired by the institution. In 1998, the Letterenhuis received a donation of De Coninck's literary archive, consisting of manuscripts and typoscripts, correspondence, diaries, notebooks, photographs and 218 floppy disks (5¼- and 3½-inch). De Coninck's paper archive is now fully [catalogued](#), but the content of the digital files is still largely unknown. The aim of the research scenario is therefore to partially ‘unlock’ the digital files stored on the floppy disks by describing their contents, creating sub-datasets that group related files while documenting their original context (e.g., the files stored on the same floppy disk) and linking the files with related files in the paper archive.

The born-digital files

The Letterenhuis created disk images of the floppy disks, which are not just copies of all the files, but a “literal representation of every bit of information on some original instance of source media”, to ensure that the original data is preserved (Kirschenbaum 2008). In total, the floppy disks contain over 1300 files (including .wpd and .doc format), which were converted by the Letterenhuis to .txt and .pdf for further processing. In some cases, the conversion rendered some characters incorrect (e.g. ‘ë’ instead of ‘é’) or completely illegible, but most of the files are perfectly suitable for further analysis. A first examination of the files revealed that they contain contributions to the Flemish literary journal *Nieuw Wereldtijdschrift* (NWT) – of which De Coninck was editor-in-chief –, minutes of editorial meetings for that same literary journal, professional *and* personal letters, and many essays written by De Coninck

himself, mostly on poetry and literature in general. For example, some of the files on the floppy disks can be linked to the following collections of essays: *De flaptekstlezer* (1992), *Intimiteit onder de melkweg* (1994) and *De vliegende keeper* (1995). Herman de Coninck also wrote most of these essays for the Flemish newspaper *De Morgen*.

Aggregating a specific sub-dataset

This presentation reports on the creation of a sub-dataset of files related to essays published in the essay collection *De vliegende keeper*, which would allow for further research, for example within the field of textual scholarship and genetic criticism. Within genetic criticism, the study of the writing process is usually divided into three levels: the examination of the exogenesis, the endogenesis (Debray-Genette 1977) and the epigenesis (Van Hulle 2014). Epigenesis describes the continuation of the genesis after a text has been published and studies the revisions in later publications (Van Hulle 2014). The files on the floppy disks can therefore be well used for epigenetic research, focusing on questions such as: What can the files on the disks tell us about the publication of this collection? Do the files tell us anything about the selection process? Did the essays need to be rewritten to be republished, or could they simply be published as they were? And what is the status of the digital files on the disks compared to the newspaper version, the version published in the collection of essays, and the typoscripts in the paper archive? As a first step in the creation of the sub-dataset dedicated to answering such questions, the (titles of the) files will be compared with the titles listed in the file entitled KEEPER (C33941/74 961), in which De Coninck listed 56 texts that could become part of *De vliegende keeper* – as well as with the 30 essays that made it into print. It will then be checked whether these essays were published in *De Morgen* and whether related files (e.g. typoscripts) can be found in De Coninck's paper archive. If so, these paper versions will be digitised.

Enriching and analysing the files

Once all the different versions of the essays have been collected and digitised, they are enriched and analysed using Diff Annotator. This is a lightweight environment for annotating text comparisons of two plain text files, developed by Vincent Neyt for the [eXtant Toolkit for Digital Scholarly Editing](#) (University of Antwerp/CLARIAH-VL), which will produce a visualisation of the variation based on Git Diff. Diff Annotator then allows for correction and annotation of the text comparison. For sustainability, the resulting HTML files could eventually be transformed into TEI-XML, to be included in the sub-dataset.

Sustaining the sub-dataset with metadata

The Collections as Data workflow will be considered in the curation of the sub-dataset (Candela, Chambers & Irollo 2023). The Collections as Data initiative encourages cultural heritage professionals to consider how they provide access to their collections to facilitate analysis using digital tools and methods. The steps in the Collections as Data workflow will be balanced against the Letterenhuis's own description model and workflow to find an appropriate way to publish the dataset and make this born-digital archival material easier to find for researchers. This will include providing public access to the metadata (including the occurrence of variance and its original context), while the dataset itself (respecting copyright) will eventually be made available with restricted access via the Letterenhuis catalogue.

References

- Debray-Genette, R. "Génétique et poétique: esquisse de méthode". *Littérature* 28.4 (1977): 19–39. <https://doi.org/10.3406/litt.1977.2073>
- Gustavo C, S. Chambers, and A. Irollo. A workflow to publish Collections as Data: the case of Cultural Heritage data spaces. (2023) Retrieved Jan 24, 2024 from <https://marketplace.sshopencloud.eu/workflow/13JvP6>
- Jaillant, L. "How can we make born-digital and digitised archives more accessible? Identifying obstacles and solutions." *Arch Sci* 22 (2022): 417-36. <https://doi.org/10.1007/s10502-022-09390-7>
- Kirschenbaum, M.G. *Mechanisms: New Media and the Forensic Imagination*. MIT Press, 2008.
- Kirschenbaum, M.G., E. Farr, K. Kraus, N. Nelson, C.S. Peters, and G. Redwine & D. Reside. "Digital Materiality: Preserving Access to Computers as Complete Environments." *iPRES 2009: the Sixth International Conference on Preservation of Digital Objects* (2009). <https://escholarship.org/uc/item/7d3465vg>
- Reside, D. "File Not Found: Rarity in an Age of Digital Plenty". *RBM* 15.1 (2014): 68-74. <https://doi.org/10.5860/rbm.15.1.416>
- Ries, T., & G. Palkó, "Born-digital archives." *Int J Digit Humanities* 1 (2019): 1-11. <https://doi.org/10.1007/s42803-019-00011-x>