Variants

The Journal of the European Society for Textual Scholarship

15-16 | 2021

Textual Scholarship in the Twenty-First Century

The Reconstruction of the Author's Movement Through the Text, or How to Encode Keystroke Logged Writing Processes in TEI-XML

Lamyk Bekius

Electronic version

URL: <https://journals.openedition.org/variants/1245>

DOI: 10.4000/variants.1245

ISSN: 1879-6095

Publisher

European Society for Textual Scholarship

Printed version

Date of publication: 1 July 2021

Number of pages: 3-43

ISSN: 1573-3084

Electronic reference

Lamyk Bekius, "The Reconstruction of the Author's Movement Through the Text, or How to Encode Keystroke Logged Writing Processes in TEI-XML", *Variants* [Online], 15-16 | 2021, Online since 01 July 2021, connection on 16 July 2021. URL: <http://journals.openedition.org/variants/1245> ; DOI: <https://doi.org/10.4000/variants.1245>

The authors

The Reconstruction of the Author’s Movement Through the Text, or How to Encode Keystroke Logged Writing Processes in TEI-XML

Lamyk Bekius

Abstract: This essay demonstrates how the use of the keystroke logging tool Inputlog allows for a fine-grained analysis of literary writing processes. But before the writing process can be studied, the keystroke logging data needs to be transformed into an output that is suitable for a textual genetic analysis. For this purpose, this essay investigates the potential of combining text with keystroke logging data in TEI-conformant XML. Besides discussing how revisions can be specified in the encoding, the author asks herself how traces of *digital* writing processes differ from *analogue* traces (and, taking it one step further, how keystroke logging can be used to record more details about the genesis of a text), what kind of decisions need to be made when encoding keystroke logging data, and how the peculiarities of digital authorship leave their mark on its encoding — as well as on the interpretation and argumentation that underlies the transcription. This will demonstrate that the level of detail that is recorded in keystroke logging data requires us to consider the way in which the text was typed when we design our encoding schemas. The goal of the TEI-XML encoding of the keystroke logging data is to provide transcriptions of writing processes that could be used to analyse (the sequence of) revisions and text production in each logged writing session in relation to their specific location in the text.

1. Introduction

THIS ESSAY ILLUSTRATES how TEI-XML encoding of keystroke logging data can be used to arrive at a detailed examination of the genesis of present-day literary works that are recorded with a keystroke logger, using the writing processes of the Flemish novelist Gie Bogaert (1958) as a case study. Bogaert divides his writing process into two stages. The first part — the “creative process” as Bogaert calls it — consists of making notes in a paper notebook. Here, he comes up with the concept and structure of the novel, writes character descriptions, and collects additional material. The writing of this notebook must be completed before he can start the second part of the writing process: the “linguistic creative process”.¹ This part consists of the actual writing of the novel, for which he

¹ This is how Bogaert, in conversation, described his own working method. For a video for the Belgian publishing house Standaard Uitgeverij in which Bogaert describes his creative writing method, see Bogaert 2013.

uses a word processor. As such, Bogaert works both in a “traditional” analogue writing environment, and in a digital one. This way, his working method aptly illustrates the challenges textual scholars and genetic critics face in the (early) twenty-first century. For decades now, geneticists and editors have had methods and tools at hand to study the traces of the writing process in the paper notebook: different colours of ink, interlinear additions, and crossed out words provide valuable clues to gain an insight into the text’s genesis. But how do we analyse the genesis of a text that was written in a word processor?

Like Bogaert, most present-day authors predominantly work in a digital environment (Buschenhenke 2016; Kirschenbaum 2016; Kirschenbaum and Reside 2013; Ries 2018; Van Hulle 2014; Vauthier 2016). Common word processors tend to hide the author’s writing operations, which makes a detailed reconstruction of the writing process that includes immediate revisions a difficult endeavour (Mathijssen 2009). Without a doubt, as Matthew Kirschenbaum and Doug Reside argue in their essay on “Tracking the Changes”, the analysis of these “new textual forms require new work habits, new training, new tools, new practices, and new instincts” (Kirschenbaum and Reside 2013, 272). The consequences of the digital work process for genetic criticism are addressed in the interdisciplinary project *Track Changes: Textual Scholarship and the Challenge of Digital Literary Writing* that combines aspects of cognitive writing process research with textual scholarship.²

Within cognitive writing process research, tools have been developed to analyse non-literary writing processes, amongst them what is called *keystroke logging*: “as an observational tool, keystroke logging offers the opportunity to capture details of the activity of writing, not only for the purpose of the linguistic, textual and cognitive study of writing, but also for broader applications concerning the development of language learning, literacy and language pedagogy” (Miller and Sullivan 2006, 1). With the aim of broadening research coverage from short professional writing processes in educational and corporate contexts to include long-term literary writing processes, the team behind the keystroke logging tool Inputlog (Leijten and Van Waes 2013) at the University of Antwerp collaborated with Bogaert to log the writing process of his tenth novel, *Roosevelt* (2016). After Bogaert’s writing process had been recorded, collaboration was established between the Inputlog team, the Literary Department at Huygens ING (Amsterdam), and the Centre for Manuscript Genetics (University of Antwerp) in order to adequately address the analysis of this literary writing process. This resulted in the interdisciplinary project *Track Changes*, in which Bogaert’s writing process is examined from the perspective of cognitive writing process research, as well as

² This work has been conducted as part of my PhD research within the project *Track Changes: Textual Scholarship and the Challenge of Digital Literary Writing* (2018–2023), a collaboration between Huygens ING (Royal Netherlands Academy of Arts and Sciences, Amsterdam) and the University of Antwerp (Antwerp Centre for Digital Humanities and Literary Criticism) funded by the Dutch Research Council (NWO). Project members include Prof. Karina van Dalen-Oskam, Prof. Dirk Van Hulle, Prof. Luuk Van Waes, Dr Mariëlle Leijten, Vincent Neyt and Floor Buschenhenke.

from the perspective of genetic criticism. In turn, the collaboration with Bogaert illustrates the possibilities of keystroke logging for genetic criticism in future collaborative projects, or when writers choose to log their writing processes themselves as part of their personal archive.³

The Inputlog team invited Bogaert to record his writing process with Inputlog. For the purpose of this essay, it is important to keep in mind that this implies that Bogaert's writing process contained a researcher intervention. Throughout the writing process itself, however, this intervention was kept to a minimum: since Inputlog is designed to be as non-intrusive as possible, the further course of the writing process was not influenced (Leijten and Van Waes 2013). Whilst other keystroke logging tools (e.g. Scriptlog, see Wengelin et al. 2009) were designed for experimental word-processing environments — and so cannot be used to study writing in a more naturalistic setting — Inputlog logs the data of writing processes that take place in a word processing environment that the author is already familiar with: Microsoft Word (Leijten and Van Waes 2013).⁴ Each time an author activates Inputlog to start a new writing session, the Word document in which the author is working is saved in the background, in a folder that contains the session's date and number. Subsequently, the Word document is saved again when the author ends the writing session by de-activating Inputlog. This results in a session-version of the text for each session, which shows the text's gradual expansion. But Inputlog does not just save Word documents. When the program is running, every keystroke and mouse movement is recorded with a timestamp (Leijten and Van Waes 2013). While writing, authors retain control of the process: they can start and stop the logging when they choose, and the data is stored on their local PC or laptop.

Although Inputlog is developed for textual and cognitive study of writing, the data output from the writing process of Bogaert's *Roosevelt*, generated in Inputlog, is not immediately suitable for literary textual research. While Inputlog provides a video replay of the recorded writing session, some issues emerge when replaying Bogaert's writing process. Short writing sessions comprising linear text production are replayed accurately, but as soon as larger segments of text are relocated or deleted, when the writing is characterized with non-linearity, or when the logged session is of considerable length, the replay mode is affected and represents the revisions and text production at the wrong location in the text. Moreover, relying solely on a video replay of the writing session for text genetic analysis also seems undesirable, as one would need to watch a writing session of, say, two hours in its entirety, while constantly pausing to analyse the effect of the revisions. A static reconstruction of the writing session — whether or not in combination with a video replay, as in Dirk Van Hulle's

³ Part of the *Track Changes* project is the development of a new keystroke logger based on Inputlog, which improves the usability and the convenience for the authors so that it can be used for their own archival practices.

⁴ For an overview of logging tools, see Van Waes et al. 2011 and Lindgren et al. 2019a.

proposal for a “Dynamic Facsimile” in the present issue (Van Hulle 2021) — is favoured to ensure adequate analysis. Hence, in order to be able to study the revisions (contained in the keystroke logging data) in their textual context, the twofold output of Inputlog — the Word document and the keystroke logging data — requires some reassembly.⁵

Since TEI-conformant XML is widely used to create a digital form of humanities data — texts, manuscripts, archival documents and so on — I opted to encode the keystroke logging data in TEI-XML to visualize and analyse revisions in their textual context (Burnard 2014). For me, these transcriptions function as a tool to gain more insight into the textual genesis. They could eventually be used for visualizations of the writing process, but a proper discussion of the latter lies outside the scope of this article. In order to reflect on how keystroke logging data can be encoded in TEI-conformant XML, this essay discusses a) the way in which the traces of *digital* writing processes differ from traces of *analogue* writing processes (and how keystroke logging can be used to record more details about the genesis of a text); b) which decisions we need to consider when we encode keystroke logging data; c) the way in which the peculiarities of digital writing leave their mark on the encoding; and d) the interpretation and argumentation underlying the transcription. The goal of the TEI-XML encoding of the keystroke logging data is to provide transcriptions of the writing processes that could be used to analyse (the sequence of) the revisions and text production in each logged writing session in their location in the text.

2. Born-digital literature and genetic criticism

The digital environment in which present-day literature is composed significantly changes the materiality of the sources available for textual scholarship and genetic criticism. Miriam O’Kane Mara argues that “with digital manuscripts, scholars must investigate the methods that authors use as they save their work

⁵ Within cognitive writing process research, a method has been developed to study revisions in context: the S-notation (Kollberg 1998). This represents the changes in the text at their location and provides information about the range, order and structure of the revisions (Kollberg and Eklundh 2002, 91). This computer-based notation can be generated using the keystroke logging data from Inputlog and is provided within the “analyse” feature of the software. However, the S-notation was initially developed to visualize revisions of short writing processes in experimental settings. As such, it appeared to be unsuitable for the study of longitudinal literary writing processes logged in their natural setting. Literary writing processes may take up several years and hundreds of writing sessions, with the production of an extensive number of words. As a result, the S-notation could not be generated using the keystroke data gathered from Bogaert’s writing process. More generally, the S-notation does not allow for further annotation and processing. Another problem concerns the representation of deleted text. Since Inputlog logs the position of the event according to its position on the x- and y-axes of the MS Word document, the deleted text is not always presented correctly (the only information the keyboard provides about a deletion is usages of the delete or backspace key). This hinders an automatically generated visualization of the revisions in their textual context.

as well as the software and hardware systems through which they compose" (Mara 2013, 345). Several explorations of the digital literary writing process have already been published since then that explore digital files, file formats, and media types.

In *Track Changes. A Literary History of Text Processing* (2016), for example, Matthew Kirschenbaum describes the emergence of the word processor and its adoption among Anglo-American authors. And in "The rationale of the born-digital dossier génétique", Thorsten Ries analysed born-digital records from the hard drives of the German poet Thomas Kling (Ries 2018; see also Ries 2017). Ries argues that the digital forensic record of the writing process comprises "digital documents", but also metadata, automatically saved draft snapshots, recoverable temporary files and other fragmented traces scattered across the hard drive (Ries 2018, 417). The use of applications such as a hex editor, a binary parser, an undelete tool, or a file carver can reveal revision and intermediate steps in the writing process (Ries 2018, 418). For example, the so-called "scratch file" [~WRS0003.tmp] with the first paragraph of the first chapter of Kling's essay *Herodot* (2005) "contains an almost complete protocol of the first writing phase of this paragraph in the form of text additions and textual variants from the first written line on to a point of time between [~WRL3681.tmp] and [~WRA1775.wbk]" — two other temporal and backup files (Ries 2017, 141). After extracting the fragments, Ries could reconstruct the nonlinear development process of this paragraph, with inclusion of editing phases and correction of typing errors (Ries 2017, 142). Still, as Ries mentions, "although the relative, layered sequence of edits can be determined [...] due to textual fragmentation, it is not in all cases possible to determine a consistent text status at any given time with certainty" (Ries 2017, 142). Bénédicte Vauthier took on the task of investigating the digital files the Spanish writer Robert Juan-Cantavella saved during the writing of his novel *El Dorado* (2008). After comparing digital documents and analysing the tree structure of the folders and other metadata such as the file title and creation date, she concluded that, "although the dossier does not contain the normal traces of writing — cancellations, additions, shifts — whose absence [...] would appear to make our analysis practically impossible, collating and comparing the digital documents and files gives us more than a sound basis to allow a meaningful genetic investigation" (Vauthier 2016, 175). Although the research above has proven that the digital writing process can leave sufficient traces to ensure genetic analysis, immediate revision and correction of typing errors remain often — apart from some exceptions — irretrievable.

Most genetic studies of born-digital writing processes work with self-archived born-digital materials received directly from the authors in question (see, for example, Crombez and Cassiers 2017; Vásári 2019; Vauthier 2016). This already indicates how important it is to collaborate with authors for this kind of research. Collaboration between writers, scholars and archivists has already been advocated by Catherine Hobbs with regard to born-digital archives, "to understand the relationship between writers, their documentation and their creative vision"

(Hobbs, qtd. in Gooding et al. 2019, 384). To be able to study the genesis of a born-digital work of literature in more detail, the collaboration with authors may be extended, for example, by logging their writing process with a keystroke logger. Between 2014 and 2018, the English novelist C. M. Taylor collaborated with the British Library to record the writing process of his novel *Staying On* (2018) using the keystroke logging software Spector Pro. Taylor, driven by both the “lost drafts” in digital writing and the loneliness of the writing process, contacted the digital curation team of the British Library, and decided together to record the writing process (Taylor 2018, n.p.). As for copyright, they agreed that the recorded data would belong to the British Library, while that of the resulting book would belong to Taylor. On the “English and Drama blog” of the British Library, Taylor quoted Jonathan Pledge (curator of contemporary archives at the British Library) who stated that the used software “seem[ed] to have been specifically designed for low-level company surveillance of employees, potentially without their knowledge” (Taylor 2018, n.p.). This use case scenario becomes apparent in the material that was recorded: the document with the text was not automatically saved after each writing session, although the software did save a screenshot “captured every few seconds each time activity on the host computer is detected”, and Taylor saved some intermediate versions himself (Taylor 2018, n.p.). Also, the software did not record the location of the keystroke within the document itself, nor the time of each individual keystroke. This suggests that while the recorded material undeniably contains valuable information about Taylor’s writing process, the software that was used seems to make it even more difficult to pursue a detailed reconstruction of the revisions in their textual context than is the case with Inputlog.

Although keystroke loggers originated as spyware, both Inputlog and Scriptlog (another keystroke logging program) have been developed specifically to observe writing processes for research purposes (Wengelin et al. 2009). The keystroke logging files logged with Scriptlog can be read and processed within the Inputlog environment (Van Hoorenbeeck et al. 2015).⁶ As a result, the proposed encoding below could also be applied to writing processes logged with Scriptlog. In addition, if it is possible to represent the revisions made by Taylor — as well as revisions logged with other keystroke loggers — in their textual context, the revision types given below may also be distinguished in this material.

3. Keystroke logging data and its use for genetic criticism

So what does Inputlog’s recorded keystroke logging material consist of? Bogaert wrote *Roosevelt* in 266 days — from July 2013 to December 2015 — during which

⁶ The same applies to eye and handwriting observation software EyeWrite, EyePen and HandSpy (Van Hoorenbeeck et al. 2015).

447 writing sessions took place, each logged with Inputlog. Of these sessions, 422 were dedicated to the writing of the novel. Luuk Van Waes assisted Bogaert in installing Inputlog, and was available for questions or when problems occurred. The 422 writing sessions resulted in 453 session-versions⁷ showing the gradual expansion of the text, and 277 hours, 14 minutes and 22 seconds of keystroke logging data.⁸

Let us zoom in on the the composition of a single sentence written by Bogaert in session thirty:

Soms kan hij meer krijgen dan wat hij voor zo'n kunstwerkje vraagt, maar dat wil hij nooit.

[Sometimes he can get more than he asks for such a work of art, but he never wants that.]⁹

All the writing actions performed in composing this sentence are listed in Table 1. Since all the different steps Bogaert took to arrive at this sentence are recorded in Inputlog's General Analysis logs, keystroke logging facilitates an analysis of the text with a granularity that cannot be obtained when working with analogue materials.

In the typology of writing processes of literary works, genetic criticism distinguishes between exogenetic, endogenetic, and epigenetic writing processes — all of which can be studied from a microgenetic point of view, and from a microgenetic one (Biasi 1996; Van Hulle 2016). "Microgenesis" here includes all intra-textual processes: "the processing of a particular exogenetic source text; the revision history of one specific textual instance across endogenetic and/or epigenetic versions; the 'réécritures' or revisions within one single version" (Van Hulle 2016, 50). The "macrogenesis" on the other hand, embodies "the genesis of the work in its entirety across multiple versions" (Van Hulle 2016, 50). When examining keystroke logging data, the geneticist can examine the writing process at an unprecedented level of granularity. The fine-grained data of keystroke logging therefore allows for a new type of what can be called "nanogenetic" research.¹⁰

Central to a work's nanogenesis would be the author's movement through the text, as the writing process is taking place. Thanks to keystroke logging software, this highly detailed form of sequentiality can be deduced from logged events that allow us to reconstruct the order in which the text was typed — for

⁷ In some sessions the initial document did not correspond with the end document from the previous session, so there are more session-versions than sessions.

⁸ For an example of how the successive events in these sessions were logged, see Table 8 in the Appendix below, which shows a detail from the General Analysis generated by Inputlog. The General Analysis of each writing session represents every event that was recorded during that session.

⁹ All the translations of Bogaert's sentences are my own. I have tried to stay as close as possible to the original Dutch sentences in my translation.

¹⁰ See also Dirk Van Hulle's contribution on "Dynamic Facsimiles" in the present issue of *Variants*.

Table 1: Writing actions in the composition of the sentence “Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij nooit.”*

<i>Writing action</i>	<i>Text</i>
adds: Soms beiden	<i>Soms beiden</i>
adds: krijgt h	Soms <i>krijgt</i> hbeiden
adds kan hij meer	Soms <i>kan hij meer</i> krijgt hbeiden
adds: en dan wat hij voor zon’ kunstwerkje vraagt	Soms kan hij meer <i>krijgen dan wat hij voor zon’ kunstwerkje vraagt</i> hbeiden
deletes: n	Soms kan hij meer krijgen dan wat hij voor zo’ kunstwerkje vraagt hbeiden
adds: n	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt hbeiden
deletes: t hbeiden	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt
adds: , maar dat il hij niet	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, <i>maar dat il hij niet</i>
adds: w	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat <i>wil</i> hij niet
adds: .	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij niet.
adds: ooit	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij <i>nooit</i> iet.
deletes: iet	Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij nooit.

example, whether the author left a sentence while composing it — and the way in which words were deleted. This level of detail of movement through a text cannot be deduced from an analogue document in which “the documentary evidence is often so complex that it becomes impossible to determine the order in which these revisions were made with any degree of certainty” (Dillen 2015, 90). The fact that the analysis of keystroke logging data here allows us to succeed exactly where traces of analogue writing processes crucially fall short warrants

* These writing actions are deduced from Inputlog’s General Analysis logs. A copy of the original logs can be found in Appendix B below.

the coinage of a new concept such as that of nanogenesis. In addition, it is important to realize that a work's microgenesis and nanogenesis can be studied separately: one could, for example, study the variation of one specific paragraph without taking the fine-grained details about the exact order of the keystrokes into account, or, conversely, solely focus on the author's movement through the text.

In the case of digital documents, Mara claims that "different types of software mould and shape the writing process, making software a collaborator of sorts" (Mara 2013, 344). Although primarily relying on how the Irish writer Nuala O'Faolain describes writing in a word processor in her memoirs, Mara indicates changes in the author's writing process as they move to the digital environment, "a process that differs from traditional drafting and revision" (Mara 2013, 345). For O'Faolain, "the digital environment provides a sense of freedom and lack of fear because so-called mistakes can be easily rectified and revision can be immediate" (Mara 2013, 346). Writing on a laptop computer provided O'Faolain with a freedom and fluidity that "indicates *for her* a willingness to play with words and structures that other media might not promote" (Mara 2013, 344; emphasis in original). It is exactly this new mode of writing that is recorded with keystroke logging, and could be described by means of nanogenetic research.

The disadvantage of keystroke logging data such as those recorded by Input-log, however, is that the processes are recorded in such detail that the logs become almost incomprehensible to the untrained eye. To make these logs more accessible to researchers, they need to be presented in a way that captures only the relevant information, and conveys the researcher's interpretation of the data in a format that is easy to read and preferably familiar to their peers. This is exactly the strength of the Text Encoding Initiative (TEI), whose guidelines recommend the use of XML tags to both transcribe the text as it is recorded on the document, and to encode the researcher's interpretation of that record in a human and computer readable format.

Within the vast realm of possibilities that the TEI allows for, we will do well to check out the work done by its Work Group on Genetic Editions (part of the TEI's Special Interest Group for Manuscripts) that produced the "Encoding Model for Genetic Editions" (2010) to facilitate the encoding of genetic phenomena. The Workgroup focuses on two main requirements: "the ability to encode features of the document rather than those of the text, and the ability to encode time, sequentiality and writing stages in the transcription" (Dillen 2015, 81). These two requirements make for a perfect starting point for a discussion of the differences between analogue and digital textual genetic material.

First of all, in case of digital writing, the spatial organization of the document no longer contains the majority of the information about the genesis of the text — given that an analogue document is a solid (static, physical) information carrier while the digital document is ephemeral (dynamic, virtual). In analogue documents "the dialectic between a document's physical limitations (as a two-dimensional surface of limited size) and the internal structure of its different

writing zones on that surface often contains important clues in the investigation of the text's writing process" (Dillen 2015, 71). For example, the position of the two text zones in the top margin of the notebook page in Figure 1 (within the black squares) and the cramped position of the word "heeft" might indicate that the text in the right-hand zone was written before the text in the left-hand one. The dynamic visualization environment of a word processor, on the other hand, discards this type of information, since it allows text to be inserted at any given position. In case of keystroke logging, the information about the text's writing process is saved in a separate file outside of the main document.

Secondly, the keystroke logging data offers detailed information about the time and sequentiality of all the activities during writing. The genetic scholar can only hypothesize about the sequence in which the two zones in the top margin of the notebook page were written, and can in no way be entirely certain about the sequence of their writing in relation to the rest of the text. When the writing process is logged, the sequence of the writing of the text can be deduced from the keystroke logging data — effectively eliminating the need for this type of guesswork.

4. What does the document reveal?

Since the "Documentary Turn" in the late twentieth century, the document has gained a special position in textual criticism — especially within the genetic orientation (Dillen 2015, 81). For our understanding of the text, and to understand its genesis, the documentary context is regarded as a crucial source of information (Dillen 2015; Pierazzo and Stokes 2011). But for born-digital materials, the spatial organization of the document becomes less substantial in the analysis of the text's genesis, since digital documents are essentially distinct from analogue ones. As Mats Dahlström has noted, digital documents cannot be defined materially. In digital documents, works are constituted "by the pattern of signals and tensions at the binary level of the material carrier" (Dahlström 2000, n.p.). Indeed, the graphical user interface (GUI), like the print layout in MS Word, only creates the illusion of the materiality of digital documents (Van Hulle 2019, 468). Yet, at the same time, as Kirschenbaum notes, "a digital environment is an abstract projection supported and sustained by its capacity to propagate the illusion [...] of immaterial behavior: identification without ambiguity, transmission without loss, repetition without originality" (Kirschenbaum 2008, 11). Indeed, as Katherine Hayles argued in "Translating Media: Why We Should Rethink Textuality", digital documents are always bound to a material carrier, in which "data files, programs that call and process the files, hardware functionalities that interpret or compile the programs, and so on" are required to produce the digital document (Hayles 2003, 274). Still, as Ries reminds us, we have to keep in mind that these digital documents "are not bound to a single physical entity, not even to a single processing system context or display application" (Ries 2018, 397). The everyday use of the term

“document” also seems to complicate its use in textual scholarship: we “speak of the ‘same’ digital document when we save ‘it’ after changing its content, after copying ‘it’ to a pendrive and ‘open it’ on a different computer with a different word processor which might display the content in a different way” (Ries 2018, 397). In my comparison with the analogue document, I therefore use the term “document” to refer to *each* Word file in the material recorded with Inputlog: each writing session creates two documents (Word files with .docx extension) and one XML file with the keystroke information (.idfx extension).¹¹ Each Word file represents a different stage of the text (session-version) and can be collated as such to gain a first impression of the development of the text during each session.

At the document level, there are two main differences between analogue source materials and born-digital source materials logged with a keystroke logger: 1) keystroke logging abolishes the distinction between interdocumentary and intradocumentary variation; and 2) the document layout no longer contains the primary information about the genesis of the text.

4.1. Interdocumentary versus intradocumentary variation

The first difference arises at the level of intradocumentary and interdocumentary variation. In writing on paper, the document pages bear witness of textual genesis through *intradocumentary* variation (Schäuble and Gabler 2018, 165). Besides this stratification visible on the page, a large part of the textual development also happens *off the page*, for example in the rewriting of a text (Schäuble and Gabler 2018, 165). This kind of *interdocumentary* variation can only be made apparent by means of collation. To account for these differences, Schäuble and Gabler proposed a distinction between textual *layers* and *levels*. Textual layers represent the intradocumentary variation (i.e. the revisions made to a single document) while textual levels describe the interdocumentary variation (i.e. the differences between two documents) (Schäuble and Gabler 2018, 169).

Encoding interdocumentary changes causes a number of interpretative problems for the editor. When there is no materialization of the change, rules need to be developed for the encoding of the revision (Schäuble and Gabler 2018, 171). To illustrate this need, Schäuble and Gabler discuss how Virginia Woolf changed the phrase “‘my mothers [*sic*] name’ (Woolf MS.A.5.b, n3)” to “‘my mothers [*sic*] laughing nickname’ (Woolf TS.A.5.a, 54)” during her transcription of her manuscript into a typescript (Schäuble and Gabler 2018, 171). This change could be encoded in several ways:

It could be encoded as a single substitution of the word “name” with “laughing nickname” or as an addition of the word “laughing” followed by a substitution of “name” with “nickname”. If we tokenise on a finer level of granularity than the word, it could even be encoded as a single

¹¹ How to preserve those files is yet another question (see, for example, Kirschenbaum et al. 2010).

addition of the string “laughing nick” that builds a new compound with the following invariant string “name”.

(Schäuble and Gabler 2018, 171)

Each of these solutions provides a correct representation of the typescript, but for the editor it is difficult to decide which encoding models the writing act best (Schäuble and Gabler 2018, 171).

For digital writing processes that are tracked with keystroke logging software, the encoding of the writing act would be less ambiguous, since interdocumentary variation is always preceded by intradocumentary variation — which (in most cases) will be recorded through keystroke logging, and saved in a separate file.¹² When using Inputlog, for example, the author can continue to write in a single document for the duration of their entire writing process, while intermediate versions are simultaneously saved as separate documents in the background. If these are logged consistently and without errors, the keystroke logging data encompasses all intradocumentary variation, which in turn provides the information about the interdocumentary variation; in this case, the difference between the Word documents saved at the beginning and the end of a writing session can be visualized by means of collation. An example from Bogaert’s writing process may help to clarify how this would work.

A collation of the Word documents that were saved at the start and the end of the seventeenth writing session points at an interdocumentary variant. Here, Bogaert changed the sentence “Mijn oude huis?” [*My old house?*] into “Mijn oude huid?” [*My old skin?*]. Following the reasoning used by Gabler and Schäuble, this could be encoded as a substitution of “huis” with “huid” or, at an even finer level, as a substitution of “s” with “d” — both options would be feasible. Since this substitution was logged with Inputlog, the question of how to change occurred is no longer an object of speculation. The keystroke logging data details the sequence of how the revision was carried out: Bogaert first pressed the key “d” and then used the delete key to remove the “s” (see Table 2).

¹² The exceptions to this are the writing sessions in which Bogaert inserted fragments of texts, which he had written in Evernote at times when he did not have his laptop at his disposal. Of these fragments, the keystrokes were not recorded.

Table 2: The replacement of “s” by “d” in the keystroke data.[†]

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock
21	keyboard	LEFT	1838	4501	4501	42308	00:00:42.308	42355	00:00:42.355
22	keyboard	d	1838	4501	4501	43228	00:00:43.228	43306	00:00:43.306
23	keyboard	DELETE	2285	4501	4502	43369	00:00:43.369	43431	00:00:43.431
24	replacement	[1839:1840]	2285	4501	4502	43369	00:00:43.369	43431	00:00:43.431
26	mouse	Movement	2285	4501	4502	45397	00:00:45.397	46083	00:00:46.083

Table 3: General Analysis of typing “Of is het m” and deleting “M”.[†]

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock
2460	keyboard	O	1862	4573	4774	1077967	00:17:57.967	1078014	00:17:58.014
2461	keyboard	f	1863	4574	4775	1078295	00:17:58.295	1078342	00:17:58.342
2462	keyboard	SPACE	1864	4575	4776	1078498	00:17:58.498	1078560	00:17:58.560
2463	keyboard	i	1865	4576	4777	1078747	00:17:58.747	1078778	00:17:58.778
2464	keyboard	s	1866	4577	4778	1078825	00:17:58.825	1078903	00:17:58.903
2465	keyboard	SPACE	1867	4578	4779	1078966	00:17:58.966	1079028	00:17:59.028
2466	keyboard	h	1868	4579	4780	1079184	00:17:59.184	1079231	00:17:59.231
2467	keyboard	e	1869	4580	4781	1079324	00:17:59.324	1079387	00:17:59.387
2468	keyboard	t	1870	4581	4782	1079418	00:17:59.418	1079465	00:17:59.465
2469	keyboard	SPACE	1871	4582	4783	1079730	00:17:59.730	1079777	00:17:59.777
2470	keyboard	m	1872	4583	4784	1080666	00:18:00.666	1080728	00:18:00.728
2471	keyboard	DELETE	1911	4583	4785	1081243	00:18:01.243	1081321	00:18:01.321
2472	replacement	[1873:1874]	1911	4583	4785	1081243	00:18:01.243	1081321	00:18:01.321

[†] Unfortunately, Inputlog occasionally gives the wrong position in *Pos.*, as is the case with the deletions in Table 2 and Table 3. The correct positions are given in *Output*, respectively [1839:1840] and [1873:1874] (as opposed to the position given in *Pos.*, respectively 2285 and 1911). For a more detailed explanation of what type of data is logged in which of Inputlog’s “General Output” columns, see Appendix B below.

Table 4: General Analysis of typing “M” and deleting “Of is het m”.[‡]

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock
8828	keyboard	M	1834	4513	5517	2647478	00:44:07.478	2647524	00:44:07.524
8829	keyboard	DELETE	1835	4513	5518	2648336	00:44:08.336	2648336	00:44:08.336
8830	replacement	[1835:1836]	1835	4513	5518	2648336	00:44:08.336	2648336	00:44:08.336
8832	keyboard	DELETE	1835	4512	5518	2648835	00:44:08.835	2648835	00:44:08.835
8833	replacement	[1835:1836]	1835	4512	5518	2648835	00:44:08.835	2648835	00:44:08.835
8835	keyboard	DELETE	1835	4511	5518	2648866	00:44:08.866	2648866	00:44:08.866
8836	replacement	[1835:1836]	1835	4511	5518	2648866	00:44:08.866	2648866	00:44:08.866
8838	keyboard	DELETE	1835	4510	5518	2648913	00:44:08.913	2648913	00:44:08.913
8839	replacement	[1835:1836]	1835	4510	5518	2648913	00:44:08.913	2648913	00:44:08.913
8841	keyboard	DELETE	1835	4509	5518	2648944	00:44:08.944	2648944	00:44:08.944
8842	replacement	[1835:1836]	1835	4509	5518	2648944	00:44:08.944	2648944	00:44:08.944
8844	keyboard	DELETE	1835	4508	5518	2648975	00:44:08.975	2648975	00:44:08.975
8845	replacement	[1835:1836]	1835	4508	5518	2648975	00:44:08.975	2648975	00:44:08.975
8847	keyboard	DELETE	1835	4507	5518	2649006	00:44:09.006	2649006	00:44:09.006
8848	replacement	[1835:1836]	1835	4507	5518	2649006	00:44:09.006	2649006	00:44:09.006
8850	keyboard	DELETE	1835	4506	5518	2649053	00:44:09.053	2649053	00:44:09.053
8851	replacement	[1835:1836]	1835	4506	5518	2649053	00:44:09.053	2649053	00:44:09.053
8853	keyboard	DELETE	1835	4505	5518	2649084	00:44:09.084	2649084	00:44:09.084
8854	replacement	[1835:1836]	1835	4505	5518	2649084	00:44:09.084	2649084	00:44:09.084
8856	keyboard	DELETE	1835	4504	5518	2649116	00:44:09.116	2649131	00:44:09.131
8857	replacement	[1835:1836]	1835	4504	5518	2649116	00:44:09.116	2649131	00:44:09.131
8859	keyboard	DELETE	1835	4503	5518	2649553	00:44:09.553	2649615	00:44:09.615
8860	replacement	[1835:1836]	1835	4503	5518	2649553	00:44:09.553	2649615	00:44:09.615

The keystroke data appears to contain even more information about the composition process of this sentence, in that it reveals another modification as well. Somewhat later in the writing process, Bogaert adds the clause “Of is het m” to the beginning of the sentence and deletes the capital letter “M” (See Table 3). The sentence now reads: “Of is het mijn oude huid?” [*Or is it my old skin?*]. But after a while, Bogaert returns to this sentence and changes it back to its previous variant by first adding “M” and then deleting “Of is het m” (see Table 4).¹³ Because this substitution was both performed and undone during the same session, it is not visible in the end document of this particular session (the session-version).

¹³ Note that Inputlog’s “General Analysis” only shows keystrokes and mouse movements. This explains why the deleted text is not visible in the log of Table 4, but only Bogaert’s pressing of the delete key. Still, the “General Analysis” *does* provide information about the position of the deleted characters. The letter “M” is positioned at 1834 and the subsequent letter at position 1835 is deleted eleven times. This allows us to reconstruct the deleted text.

When we try to encode this sentence, merging the text with the keystroke data, all the modifications can be put together as follows:

```
<seg><add>M</add><del><add>Of is het m</add></del>
<del>M</del>ijn oude hui<add>d</add><del>s</del>?</seg>
```

(Example 1)

The keystroke logging data thus provides more information about interdocumentary revisions than we can extract from analogue material: all interdocumentary variation is captured as intradocumentary variation within the keystroke logging data. When a keystroke logger is used while writing, textual development can no longer occur *off the page* — or at least: when authors keep their writing process within the confines of their Inputlog enabled computer, and when that process is logged without errors). In Bogaert's case, the records suggest that he occasionally pasted new or revised textual materials into the document that were produced in between explicitly logged writing sessions. This means that for these passages, there is no data available about how the variation came to be.

4.2. Encoding keystroke logging data instead of the document's layout

This leads us to the second difference between analogue and born-digital writing processes at the document level, which concerns the appearance of the document page itself. Complex handwritten draft materials often consist of chaotic pages that contain multiple textual fragments that were written at different positions and in different directions on the page. The page of Bogaert's notebook (See Figure 1) shows text written in different colours, in the margins of the page, and between two lines. If we want "to gain insight about the composition, time of revisions, and flow (flux) of the text", we therefore need to carefully consider the physical aspects of the document (its layout, the arrangement of the text on the page, and what this tells us about how the text was written), and to inform that reading of the page with a good understanding of the text (Workgroup on Genetic Editions 2010, §1.3). But when Bogaert uses a word processor, the document remains clean with every modification to the text (see Figure 1): additions are always represented as inline insertions, and deleted text "disappears" from the surface. Logging the process with a keystroke logger that saves these modifications in a separate file prevents them from becoming untraceable. The crucial information about the genesis of the text thus shifts to the keystroke logging data. This calls for a different encoding of the revisions, that focuses on exactly those elements that make a difference in the keystroke logging data.

When the text is composed in a word processor, revisions cannot be specified with the attributes used in encoding analogue material. Instead of indicating the specific writing tool (which may be encoded in @rend) or the location in the

Figure 1: A comparison of Gie Bogaert's analog and digital writing processes. Left: a page in Bogaert's notebook his novel *Roosevelt* (left). Right: a screenshot of a page in one of Bogaert's MS Word documents for the same novel.

document (which may be encoded in @place), digital revisions (specifically: <add>s and s) can be further specified using their location in the text. These diverging behaviours in the writing of analogue versus digital documents forces us to completely rethink the ontology we use for encoding relative location in our transcriptions. In the following, I therefore propose a list of “revision types” on the basis from Lindgren and Sullivan’s so-called “revision taxonomy” (2006a). By using their taxonomy as a starting point, we can encode the relative location of additions and deletions in the attributes of <add> and tags in a way that is more relevant to born-digital writing processes (see Tables 5 and 6).¹⁴

The first step in adopting Lindgren and Sullivan’s taxonomy is to define revisions according to their relative location in the text, that is “where and when in the writing process revision occurred” (Lindgren and Sullivan 2006b, 42). With regard to this location, the taxonomy distinguishes “pre-contextual” revisions (i.e. “revisions made before an externalised context is completed”) from “contextual” revisions (“revisions made within a completed externalised

¹⁴ For good measure, the attributes for additions and deletions in witnesses of born-digital writing processes (Table 6 are contrasted to those for analogue writing processes (Table 5). The specific examples that were used to compile these tables are taken specifically from analogue and digital witnesses to Bogaert’s writing process, but could be applied more generally.

context"; see Lindgren and Sullivan 2006a, 159).¹⁵ These location-based revision types best resemble the use of attributes like *@rend* or *@place*, as they are based on the relationship between the keystroke data and the place in the document, rather than purely on the editor's interpretation.

In the encoding of the keystroke data, these revision types may be applied to the textual unit of a sentence.¹⁶ Building upon the taxonomy by Lindgren and Sullivan, the attribute *@type="context"* can be used to indicate that the revision is a contextual one, that is: a revision made in a previously written sentence. The attribute *@type="pre-context"*, by contrast, may then be used to indicate revisions made before a sentence is completed and so concerns the author's most recently typed characters. Diverging from Lindgren and Sullivan's definition of pre-contextual revisions, pre-contextual deletions can take place at a point in the text with externalized text after the deleted text.¹⁷

A large number of revisions in digital writing occur as a result of typographical errors. Within the scope of genetic criticism, such "typos" are not the most meaningful entities because they do not immediately affect the meaning of the text. Within cognitive writing process research, typos are regarded as a revision type that "often blur[s] the picture of the writing session" (Kollberg 1998, 68). Typographical errors are "low-level, and hence less important, types of revision", and filtering them out would therefore allow for a more nuanced analysis of revision (Conijn et al. 2019, 71). But the revision of typographical errors can also break the flow in writing and therefore influence the writing process (Conijn et al. 2019, 72). For this reason, I propose to encode this type of revisions with a separate *@type* attribute: *@type="typo"*. This allows such errors to be filtered out in visualizations where they are irrelevant, while still allowing us to evaluate their effect on the writing process.¹⁸

¹⁵ Within cognitive writing process research, the writing process is generally divided into three consecutive components: first the text has to be planned, then these internal ideas have to be translated (externalized) into linguistic forms, and then those forms have to be evaluated and revised where necessary (Lindgren and Sullivan 2006b). While writing, different types of revision occur, and revision is mostly understood as "making any changes at any point in the writing process" (Fitzgerald 1987, 484; Lindgren et al. 2019b, 346). Two major categories for such revisions are *internal* and *external* revisions. The former encompasses "overall, conceptual revision as well as conscious evaluative revision and revision of pre-text" (Lindgren and Sullivan 2006b, 37). The latter are all "visible changes made to the written text" (Lindgren and Sullivan 2006b, 37). Inputlog only logs revisions made in already externalized text; the encoding therefore covers the external revisions.

¹⁶ In the proposed encoding scheme, each sentence is encoded with a *<seg>* tag.

¹⁷ In the taxonomy by Lindgren and Sullivan (2006), one feature of a pre-contextual revision is that there is no externalized text after the place of revision (Lindgren and Sullivan 2006a, 159). As literary writing is often a non-linear process, new context can be created in other places than at the end of the text. In order to be able to distinguish revision within a sentence before this sentence is completed, I also regard these revisions as "pre-contextual".

¹⁸ Typing errors can be hard to distinguish from spelling errors. In the encoding of

The use of a keystroke logger allows for an exact reconstruction of the textual development. This includes the moment a new sentence is produced. Therefore, the production of new sentences can also be incorporated in the encoding (`@type="nt"; "new text"`), to be able to differentiate between “new” and “old” sentences.¹⁹ Writing is not always a linear process and sentences are not always finished before modifications are performed elsewhere in the text. The author could, for example, move away from the point in the writing where new meaning is produced: the so-called “leading edge” (Lindgren et al. 2019b). In the definition formulated by Lindgren et al., the leading edge is located “typically at the end of the text produced so far, but can also occur at the end of insertions within previously written text where a writer inserts new ideas (not only revises form)” (Lindgren et al. 2019b, 347). Unlike the point of inscription, which comprises “all writers’ actions in previously written text as well as at the end of the text produced so far”, the leading edge is restricted to the creation of new meaning (Lindgren et al. 2019b, 347). During the production of a sentence the author can decide to leave the sentence produced so far to make a revision elsewhere — in the same sentence or at another segment of the text — after which they return to the end of the sentence they were writing. This would not be an addition, because the sentence is not yet completed. However, the fact that the author moved away from the leading edge is meaningful for the interpretation of the writing process, as it provides information about the steps that were taken to write the sentence. To be able to identify this return to the leading edge, the text can be encoded using `<mod>`. According to the TEI P5 Guidelines, the `<mod>` element may be used to represent “any kind of modification identified within a single document” (TEI Consortium 2020, §11.3.4.1). For the purpose of analysing digital writing processes, it may also be used for the “modification” of unfinished sentences — the continuation of writing the sentence — using the attribute: `@type="continue"`. A transcription with the inclusion of `<mod>` tries to model the flow of writing.

the typing errors, I therefore used a list of criteria (developed by Stevenson et al. 2006) for distinguishing *typing* revisions from *spelling* revisions. According to the checklist developed by Stevenson et al., a revision can be identified as a typing revision, if one or more of the following applies: “a. the pre-revision form does not conform to the orthographic rules of the language; b. the pre-revision form involves a letter string which does not conform to a likely pronunciation of the word; c. the semantic context indicates that the pre-revision form could not have been intended; d. the same word is written correctly at an earlier point in the text; e. a letter is replaced by an adjacent letter on the keyboard” (Stevenson et al. 2006, 232).

¹⁹ Some text segments may also be deleted and added again in a revised form; thereby maintaining a semantic relationship with the previously deleted text. This is not necessarily “new” text and may therefore be given another attribute: `@type="rt"` (‘revised text’). However, the encoding of such revised text adds a new level of interpretation to the transcription. Whereas “new text” is a fairly objective interpretation — as the text is typed into the document for the first time — the classification “revised text” relies on the editor’s interpretation and their understanding of the text.

Table 5: Specifications for additions and deletions in analogue material.

<add>	@hand="#GB"	@rend="#blueInk" @rend="#blackInk"	@type="#alternative"	@place="#marginleft" @place="#supralinear" @place="#inline"
	@hand="#GB"	@rend="#blackInk"	@type="#crossed Out" @type="#overwritten" @type="#instant Correction" @type="#underlined"	

Table 6: Specifications for additions and deletions in digital material.

<add>	@seq="201308230815"	@type="context" @type="pre-context" @type="typo" @type="nt" @type="translocation"	@n="15"	@evidence="6514-6556"
	@seq="201308230830"	@type="context" @type="pre-context" @type="typo" @type="translocation"	@n="16"	@evidence="6557-6567"
<mod>	@seq="201308230840"	@type="continue"	@n="17"	@evidence="6576-6599"

Most of these revision types can be illustrated using the steps taken by Bogaert when he wrote the example sentence from Table 1 (“Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij nooit”). He started by writing “Soms beiden”, then moved his cursor between the two words using the left arrow key. There he wrote “krijgt h”. This is a pre-contextual addition, because it takes place before the sentence is finished.

```
<add type="nt">Soms <add type="pre-context">kan hij
meer </add><add type="pre-context">krijg<mod
type="continue">en dan wat hij voor zo< type="typo">
n</del>'<add type="typo">n</add> kunstwerkje vraagt
</mod> <del type="pre-context">t h</del></add><del
type="pre-context">beiden</del><mod type="continue">,
maar dat <add type="typo">w</add>il hij n<add
type="context">ooit</add><del ="context">iet</del>
</mod><mod type="continue">.</mod></add>
```

(Example 2)

Bogaert then continued writing with another pre-contextual addition between “Soms” and “krijgt”: “kan hij meer”. After inserting this fragment, he relocates the cursor between the letter “g” and the letter “t” in the word “krijgt” and continues writing. Bogaert left the leading edge (the point where he created new meaning) to make the pre-contextual additions, but after these

insertions a new leading edge is created as he continues writing the sentence between the letter “g” and the letter “t”. At the new leading edge he writes: “en dan wat hij voor zon’ kunstwerkje vraagt”; the screen would now have displayed the sentence as:

Soms kan hij meer krijgen dan wat hij voor zon’ kunstwerkje vraagt
hbeiden.

The new leading edge was not positioned at the end of the unfinished sentence, but after the letter g in the word “krijgt”. It was thus followed by “t hbeiden”.

Bogaert then corrects the typo made in the production “zon’” (the apostrophe was incorrectly positioned) after which he eventually deletes the bulk of unused characters at the end of the sentence. These are all pre-contextual deletions; the sentence is still not finished. Now the leading edge is positioned at the end of the sentence, where Bogaert continues writing: “ , maar dat il hij niet”. After correcting the typo with an addition — he missed the letter “w” in writing “wil” — he types the full stop. This marks the moment the writing of the sentence is finished. Somewhat later in the session, Bogaert returns to the sentence to make a contextual revision. He substitutes “niet” with “nooit” by adding “ooit” and deleting “iet”. The writing process of this sentence illustrates the complexity of digital writing, but also demonstrates that the proposed encoding succeeds in capturing every step in the process.²⁰ Still, this encoding misses an important aspect of the writing process: time.

5. Specific encoding of time

Inputlog logs every keystroke and mouse movement in combination with a timestamp. Unlike analogue writing processes, the keystroke logging data allows us to incorporate the specific time of writing into the encoding. Through this temporal aspect, the writer can — so to speak — be followed through the text. Lindgren and Sullivan mention this aspect of keystroke logging too when they argue that the location of revisions

shows how the writers move their points of focus during text composition; this can be viewed as the route writers take through their texts. The actions writers perform during composition can, for example, hint at the writers’ developing ideas and associated shifts in text focus.

(Lindgren and Sullivan 2006b, 39)

Incorporating the recorded time of the revision into the encoding thus offers a

²⁰ Bogaert explained in conversation that, when writing in the Word document, he focuses primarily on finding the right words. For him, this is the hardest part of the writing process. The rather complex way in which he types this sentence could reflect this. While doing so, he seems primarily orientated towards its reformulation. However, this also reflects Bogaert’s personal typing habits.

unique opportunity to study the text's genesis at a microscopic level — what I referred to as its “nanogenesis” earlier.

The timestamp enables the genetic scholar to investigate the location at which the author was working before they made a revision at another place in the text, when (and how quickly) revisions were made, and if there were certain revision campaigns. To analyse this, the editor may encode the timestamp for each addition and deletion and every other event worth mentioning, by using the `@seq` attribute (e.g. `@seq="yyyymmddhhmmss"`). The editor can choose to incorporate the dates of the writing sessions, so as to visualize the chronology of the writing process, not only within a single session but also across several (or all) sessions. The hours, minutes and seconds indicate the time after the start of the session added to the time the session is started.²¹ As such, this notation provides the exact time the textual input took place. The TEI P5 Guidelines propose the attribute `@seq` (sequence) for assigning “a sequence number related to the order in which the encoded features carrying this attribute are believed to have occurred” (TEI Consortium 2020, §11.3.1.4). In the case of the logged writing processes, the `@seq` attribute can be very specific as the data provides information about the time the deletions and additions were being made.

The timestamp given in `@seq` can subsequently be used to number all the changes in `@n`. Using an XSLT script, the events can be listed chronologically in `<listChange>` and allocated a number.²² The `@n` includes the number of appearances of all insertions and deletions, as well as all returns to the leading edge. From a computational perspective, the number in `@n` provides the same information as is given in `@seq`: the chronology of the modifications. The benefit, however, is for the (human) reader. In the eventual transcription, the numbers will offer the reader the possibility to see the sequence of the revision in one glance (see section 7). This is one step into making the complexity of the (digital) writing process more easily analysable for the reader. The example below shows the encoding of the same sentence discussed above, with the inclusion of the time and the order of appearance, starting from 27.

```
<add seq="20130826151155" type="nt" evidence="1342"
n="27">Soms <add seq="20130826151216" type="pre-context"
evidence="1400-1411" n="29">kan hij meer </add><add
seq="20130826151210" type="pre-context" evidence="1374-
1383" n="28">krijg<mod seq="20130826151221"
type="continue" evidence="1423" n="30">en dan wat hij
voor zo<del seq="20130826151232" type="typo"
```

²¹ The time is derived from the “StartClock” in the “General Analysis” of Inputlog, which is added to the start time of the session in question. In order to be able to retrieve the event in keystroke data, the unique ID of each event in the keystroke data needs to be included in `@evidence`. As the time of every keystroke is given, the editor needs to make a decision as to which time to incorporate in the encoding. For a genetic analysis, the time of an event's first keystroke may be the most fitting option; for example, when the first key is pressed to start production of a new sentence.

²² I would like to thank Vincent Neyt for writing the XSLT script for this purpose.

```
evidence="1508-1509" n="31">n</del>'<add
seq="20130826151233" type="typo" evidence="1513" n="32">
n</add> kunstwerkje vraagt</mod><del seq="20130826151236"
type="pre-context" evidence="1552-1557" n="33">t h</del>
</add><del seq="20130826151237" type="pre-context"
evidence="1559-1581" n="34">beiden</del><mod
seq="20130826151241" type="continue" evidence="1595-1643"
n="35">, maar dat <add seq="20130826151248" type="typo"
evidence="1643" n="36">w</add>il hij n<add
seq="20130826154211" type="context" evidence="10622-
10625" n="170">ooit</add><del seq="20130826154212"
type="context" evidence="10626-10633" n="171">iet</del>
</mod><mod seq="20130826151254" type="continue"
evidence="1666" n="37">.</mod></add>
```

(Example 3)

The number gives the exact order in which the modifications were carried out while keeping editorial interference to a minimum. This contrasts with analogue sources, where the complexity of documentary evidence turns the numbering of revisions into a highly interpretative act resulting only in speculative readings (Dillen 2015, 90). By comparison, the keystroke data allows for a detailed reconstruction not only of the revisions made at sentence level, but those at complete-text level as well. If the author first made a revision to a sentence in the middle of the text and then another in a sentence at the top of the text, this movement through the text can be reconstructed, and also — crucially — referenced in analyses of the writing process.²³

6. Peculiarities of digital writing

In the encoding of the keystroke logging data, the way the text is typed is taken into account. This way, we can distinguish between different typing styles. In this respect, at least two characteristics in digital writing become apparent: 1) the recycling of words and characters, and 2) the different ways of performing a deletion. These characteristics may guide editors in the decisions they make in the encoding of born-digital writing processes.

6.1. Recycling

Although the act of deleting is effectively free of cost in a word processor (Sullivan 2013, 256), authors might recycle words and characters in rewriting their texts. This characteristic is also noted by Py Kollberg in her study of digital revisions (1998). She discusses how a writer in her corpus keeps the “t” in substituting “there” for “it”:

²³ The transcriptions below (see Figure 2 and Figure 4) and the discussion thereof, provide an example of how the numbers in @n can be used to quote specific revisions.

Probably in order to minimize the effort to make this change, the writer keeps the *t* in *it*, and uses it in the new word *there*. Two elementary character level revisions performed at different positions are the result, but the effect of both revisions is at the word level (and the words are at the same position). Many writers would have deleted the whole word in this situation.

(Kollberg 1998, 78; emphasis in original)

Kollberg concludes that people develop personal habits in their use of the word processor; each writer has their “own personal set of organization of operations” and is used to performing “certain actions in certain ways” (Kollberg 1998, 78). Not unlike handwriting, typing styles contain a “fingerprint” of the writer (Lindgren et al. 2019a, 5).²⁴ Because genetic criticism is interested in the author’s way of working, the way in which they make use of the word processor also needs to be apparent in the transcription.

As for Bogaert’s way of typing, his recycling of words is very prominent. Indeed, it is already present in the sample sentence discussed above (see Examples 2 and 3). Here, Bogaert added the clause “en dan wat hij voor zon’ kunstwerkje vraagt” between “krijg” and the letter “t” of the word “krijgt”. As such, he recycles the word part “krijg”, re-using it in the word “krijgen”. This re-use can be detected at the word level as well, as Bogaert kept the letter “n” in the substitution of “niet” with “nooit”. This is quite characteristic of Bogaert; as Kollberg remarked in a similar situation quoted above, many others would have deleted the entire word.

This recycling of words and characters makes transcription of the writing process a complex matter, as it makes the concise representation of the flow of writing more challenging. This characteristic therefore highlights the importance of encoding the returns to the leading edge with a different tag. In the genetic transcription (Figure 2), it is possible to reconstruct that “en dat wat hij voor zo’n kunstwerkje vraagt” (n30) was written between “krijg” (n28/1) and “t h” (n28/2) while taking into account that this is not a regular addition, but the writing of the sentence itself. In this visualization, the process of the writing is emphasized and the singularity of Bogaert’s writing accentuated. Hence, it is important to encode the separate steps in the writing process in order to be able to reconstruct the flow of writing.

²⁴

This also depends on the computer or laptop used; the keyboards used in a desktop PC set-up are more likely to incite use of the delete key, which is not available as a single button on many laptop keyboard layouts — see also section 6.2.

²⁷Soms kan hij meer ²²krijg ³⁰en dan wat hij voor zo ²⁴n ³²kunstwerkje vraagt t ²²h ²³beiden ³⁴35, maar dat
²⁶w ²⁶i hij nooit ¹²⁰tet ²²⁴37.
 ★ ²⁷Soms < kan hij meer < ²⁹< krijg ☆ ³⁰en dan wat hij voor zo △ n △ ²⁴△ n △ ³²kunstwerkje vraagt ☆ < 4
 h < ²²< ²³< beiden < ³⁴☆ ³⁵, maar dat △ w △ ³²il hij n > ooit > ¹²⁰> iet > ²²⁴☆ ☆ ³⁷☆ ☆

Figure 2: Transcription of a paragraph in session 30, showing all the different modifications.
 For a legend of the colours and symbols used in this transcription, see Table 7 in Appendix A below.

6.2. Deletions

Another way in which typing styles become apparent is the usage of the keyboard in making a deletion. As Kollberg notes, a delete operation can be performed in two directions: forwards and backwards (Kollberg 1998, 29). A forward deletion removes characters to the right of the cursor, a backward deletion those to its left (Kollberg 1998, 29). A forward deletion may be carried out by using the delete key or by selecting the characters to the right of the cursor and pressing the backspace key. Using only the backspace key performs a backward deletion. The way the writer uses the keyboard in performing revision affects the encoding of the revision.

When a writer uses the backspace key to delete characters in a substitution — a backward deletion — the deleted word will usually appear in front of the inserted word. If the author uses the backspace key to delete words during the production of a sentence, the cursor is continuously positioned at the end of the leading edge. During production of the clause in Example 4, “de sheerne khoran kan worden gebracht” [*the sheerne khoran can be brought*], Bogaert changed the simple past tense verb “kon” into the simple present tense “kan”. After writing “kon”, he deleted “on” and then continued writing by typing “an”:

```
<seg>[...]de sheerne khoran k<del seq="20140717142431"
type="pre-context" evidence="6689-6690">on</del>an
worden volbracht [...]</seg>
```

(Example 4)

This pre-contextual deletion can be considered as the digital equivalent of *currente calamo* deletions in analogue material, which “usually characterize writing produced by an author in the throes of composition, with corrections or revisions made immediately rather than later” (Beal 2011, 104). The linearity of the pre-contextual deletions and the production of the sentence facilitate the readability of the encoding.

The author may also use the delete key to remove a part of the text — a forward deletion. Bogaert prefers this technique: when he makes a substitution, he writes the addition prior to the deletion so that the new word appears to the left of the

older one — in the substitution of “niet” with “nooit”, for instance, the writing of “ooit” preceded the deletion of “iet”. The addition therefore appears before the deletion in the encoding. In the encoding of analogue material, however, Elli Bleeker notes that a deletion is normally located

before [i.e. “to the left of”] an addition in a transcription (regardless of the actual positioning of these elements [on the document]), simply because — in the western world — we read a transcription from left to right and we usually assume that a word is first deleted and then replaced.

(Bleeker 2015, 98)

This choice is usually guided by the goal of the transcription of analogue material: to render the text more readable (Bleeker 2015, 98). In the transcription of a digital writing process, the goal is also to reconstruct that process — as this is not visible in the document — and to capture the author's way of working. The additions and deletions are therefore best placed in the position at which they occurred: the way the deletions are performed dictates the decisions made in the encoding.

7. Interpretation, selection and argumentation

The proposed encoding produces a transcription of the keystroke logging data in order to provide data output suitable for a genetic analysis. Specifically, it allows for the examination of revisions and new text production, their sequences and their effect on the text. This transcription alone is not sufficient to create a digital genetic edition, but it provides a sound basis for the visualization of the writing process. Moreover, the act of encoding the keystroke logging data does coincide with the encoding practice for analogue material in that, here too, “relatively simple text encoding forces us to make editorial decisions” (Bleeker 2015, 112). In the case of keystroke logged writing processes, the need for abstraction (and therefore interpretation) of the recorded material only increases, because there is so much additional information available to the editor. As the examples above demonstrate, simply converting the writing actions that are recorded in the logging data to their editorial representations already implies making a series of editorial choices, such as selecting the data and deciding where the encoded insertions and deletions should be located in the transcription. While the keystroke logging data serves to make more objective observations about the sequence of the writing, it also forces the editor to make their interpretation of the material even more explicit.

The transcription of the keystroke logging data tries to be as objective as possible. In its proposal to complement a text-oriented approach with a document-oriented one, the TEI Ms SIG refers to the opposition in German editorial theory — as coined by Hans Zeller — between the “Befund” and the “Deutung”. Respectively, these refer to “what is there in the source document, the record” (Befund), and “the interpretation of this phenomenon” (Deutung) (TEI Consortium 2020,

§1.1). The Workgroup notes that one cannot talk about the record without any interpretation (especially not in the realm of genetic criticism) but does make a distinction between different levels of interpretation:

there is an obvious difference between the interpretation that some trace of ink is indeed a specific letter and the assumption that a change in one line of a manuscript must have been made at the same time as a change in another line because their effects are textually related.

(TEI Consortium 2020, §1.1)

The Workgroup therefore proposes making a distinction between the interpretation of “what’s there” (document/fact) and “how does it relate” (text/interpretation) (TEI Consortium 2020, §1.1). A similar distinction is made in research into cognitive writing process, when it differentiates between what are called *elementary revisions*, and *interpreted revisions*. According to Kollberg, an elementary revision is a single deletion or insertion, and the analysis of such elementary revisions is therefore based only on “the writer’s overt action in manipulating the text, with a minimum of interpretation of how revisions may be related according to the writer’s intentions” (Kollberg 1998, 16). Interpreted revisions, on the other hand, are revisions which are analysed at a higher level. For example, if “two or more elementary revisions that are seemingly united by the same goal may be combined and interpreted by the researcher as a unit” (Kollberg 1998, 17).

Following that logic, the proposed encoding therefore focuses solely on the elementary revisions. For example, the distinction between contextual and pre-contextual revisions rests only on the author’s actions as they are recorded in the keystroke logging data. When a revision is made during the production of a sentence (before the author presses the full stop), it is marked as a pre-contextual revision. When the revision is made within a completed sentence (after the full stop is typed), it is marked as a contextual revision. In addition, revisions are encoded according to the way they are performed and the replacement of one word with another is not encoded as a substitution.²⁵ Nevertheless, this can only allow for a certain degree of objectivity, since the selection of the material already involves interpretation (Dillen 2018, 38).

Selection plays a pivotal part in the encoding of the keystroke logging data. Although the transcription sets out to represent the author’s movement through the text, many indications of movement have been left out of the encoding. The encoding marks only the textual output, as generated by the keyboard, consisting of characters and punctuation marks. As such, it omits the keyboard events “UP”, “DOWN”, “LEFT” and “RIGHT”. The same applies to the mouse movements. In focusing on the textual output, the encoding also ignores data provided about pauses, their locations and the timing of each action. Moreover, as the only time that is encoded in the transcription is the start time of any modification, the time

²⁵ For example, when “niet” is changed to “nooit” by adding “ooit” and deleting “iet”, only “ooit” and “iet” are marked with elements.

between the end of one revision and the start of another is omitted as well. This means that the time between two subsequent revisions cannot be deduced. This does not imply, for instance, that long pauses or other time indications cannot be encoded in the transcription, but rather that such an encoding does not lie within the scope of this particular transcription. The main aim of this transcription is to help the scholar/reader follow the sequence of text production and revisions with a focus on the text and its meaning.²⁶ By reducing the presentation of other kinds of information, the scholar/reader is less distracted from analysing the text. Still, even with a focus on the textual output, interpretation remains a key factor in the encoding as “the idea of presenting a text in an objective way is problematic and arguably impossible” (Bleeker 2015, 114). This seems even more true when encoding keystroke logging data, as the editor is setting out to reconstruct a state of the text which has never existed in full.

When all the deletions and insertions within a given session are encoded, we arrive at a state of the text that has never actually appeared on the author's computer screen as such, and has therefore never interacted with. This might present us with some issues, such as the question where to encode insertions and deletions when several revisions are located at virtually the same position in the logs (Kollberg 1998, 34). The editor's interpretation is necessary in these instances, especially when the author makes an insertion next to a place in the document where there was a previous deletion. That is the case because when the author inserts text, the text is inserted at the cursor location. But since any previously deleted text remains visible in the encoding, there is no straightforward place to locate the insertion in relation to the previously deleted text (Kollberg 1998, 34). A protocol for such cases could be that when it pertains a single insertion, the insertion should be encoded to the right of the previous deletion — in line with Bleeker's argument for encoding deletions and additions in analogue witnesses. And when it comes to substitutions, the way in which the insertion and deletion were performed could help make the most accurate decision. For example, when new text was inserted first, and the old text forward deleted afterwards, we could transcribe the insertion first (i.e. to the left) and the deletion second (i.e. to the right).

The editor's interpretation also comes to the fore in the transformation of the TEI-XML encoding. Joris van Zundert and Tara Andrews argue that the interface of the digital edition functions as an argument: “Our first observation is that a digital edition's interface is an argument — not just an argument about the text, but also an argument about the “attitude” of the editor, a window into his or her take on methodology and the digital edition itself” (2018, 7). The interface of a digital scholarly edition “is always closely linked to the data model of the underlying data and the editorial principles expressed in this data model”, so they function as “an interpretation of knowledge and provide users with

²⁶ Providing another transcription in which the pauses are encoded would indeed be useful, as it would provide information about the fluency of the writing. This may help interpret the sequence of the revisions from a cognitive perspective.

a more or less “guided” tour through the data and its general presentational setting” (Bleier and Klug 2018, VII). While different transcriptions given below cannot be considered as a fully developed interface, they already function as “an integral part of rhetorical form” since they foreground the textual development (Andrews and Van Zundert 2018, 8). As an example, I shall discuss some possible transcriptions of the paragraph that includes the example sentence I referred to above, with a view to guiding attention towards the dynamics and non-linearity of the writing process — in this case within a single writing session.

The first option is to display a transcription that simply presents a reconstruction of all the textual operations within their textual context. This transcription promotes reading of the text with all the modifications made during this session. The different types of modification are visualized in different colours, which indicate that the writing of this paragraph proceeded in different steps. At a glance, one can see the dynamics that underlie the writing process:

~~Laat en ik we hem voor het gemak de krijtkunstenaar noemen.~~ Niemand weet hoe hij heet of waar hij vandaan komt. ~~Laat ik hem voor het gemak de krijtkunstenaar noemen.~~ ~~Erg~~ ~~echt~~ ~~veel~~ ~~kan ik niet over hem~~ ~~kwijt~~. Hij is hoogst bescheiden. Nogal ~~get~~steld op zijn privacy ook. Hij leeft van wat hij verdient met de ~~sch~~ilderijtjes die hij ~~maakt en verkoopt~~ ~~aan en paar vrienden en wat verwanten~~~~famileleden~~. 'Prutserij,' noemt hij ~~zijn werk~~het ~~zelf~~. Maar het is best ~~we~~l goed. Soms ~~kan hij meer krijgen dan wat hij voor zo'n~~ kunstwerkje vraagt ~~heiden~~, maar dat ~~wil hij nooit~~het. Hij is ~~ook~~ dooddeerlijk. ~~Het~~ ~~ij~~ is zo'n ~~jone~~gen waarvan ~~wie~~ je alleen maar kan houden.

Figure 3: Transcription of another paragraph in session 30, displaying all the different modifications.

The transcription can then be modified to show how the text developed during the session. By removing all the added text, this transcription visualizes the text within the paragraph that was already written at the moment Bogaert started this new session:

~~Laat en we hem voor het gemak de krijtkunstenaar noemen.~~ Niemand weet hoe hij heet of waar hij vandaan komt.

Figure 4: Transcription of the same paragraph in session 30, displaying the text as it was at the beginning of the session.

Conversely, the deleted text can also be removed. This enables readers to see the state of the text at the end of the session. By providing the option to read the state of the text at the beginning and the end of the session, one is encouraged to focus on how the text developed and on which steps were taken during its writing. This shows that the first and the second sentence were transposed and that Bogaert added eight new sentences:²⁷

²⁷ In fact, Bogaert added nine of them: one of the sentences was then deleted during the same session.

Niemand weet hoe hij heet of waar hij vandaan komt. Laat ik hem voor het gemak de krijtkunstenaar noemen. Hij is hoogst bescheiden. Nogal gesteld op zijn privacy ook. Hij leeft van wat hij verdient met de schilderijtjes die hij maakt en verkoopt aan en paar vrienden en wat verwanten. 'Prutserij,' noemt hij zijn werk. Maar het is best goed. Soms kan hij meer krijgen dan wat hij voor zo'n kunstwerkje vraagt, maar dat wil hij nooit. Hij is doodeerlijk. Het is zo'n jongen van wie je alleen maar kan houden.

Figure 5: Transcription of a paragraph in session 30, displaying the text at the end of the session.

Next, the sequence of all the modifications can be studied by displaying their numbers in sequential order. The first modifications in this paragraph are the insertions of four new sentences (n14-n17),²⁸ the writing of the fourth of these is interrupted by typing errors in “shilderijtjes [sic]” (n18-n20; [paintings]), which Bogaert corrects before finishing the sentence with “die hij maakt en verkoopt.” (n21; [which he makes and sells]). Hence, if one continues following the sequence of all the writing operations in chronological order, the nanogenesis can be analysed. This shows, for example, that Bogaert wrote the seventh sentence (n22) and then returned to the sixth sentence to insert the clause, with typos, “aan en [sic] paar vieden [sic] en famileileden [sic]” (n23; [to a few friends and family members]). He then immediately substituted “famileileden [sic]” (n25; [family members]) with “verwanten” (n24; [relatives]). The sentences “Maar het is best wel goed” (n99; [But it is pretty good]) and “Het is zo'n jongen van wie je alleen maar kan houden” (n150; [He's the kind of guy you can only love]) were inserted later, which indicates that he did not work continuously on this paragraph but relocated the point of inscription to other locations in the text in the meantime. The order of some revisions might also indicate a relationship between them; the quest to correct typos (n91-n94) and inserting a new sentence after deleting another (n149-n150). The sequence of the modifications shows the continuous shaping and reshaping of the text by Bogaert, a process made possible by the word processor:

Tag¹⁴³en¹⁴⁴ik¹⁴⁵we¹⁴⁶hem voor het gemak de krijtkunstenaar noemen.¹⁴⁶ Niemand weet hoe hij heet of waar hij vandaan komt. Laat ik hem voor het gemak de krijtkunstenaar noemen.¹²⁷ 14 Erg⁸²cht⁸³veel kan ik niet over hem kwijt.¹⁴⁹ 15 Hij is hoogst bescheiden. 16 Nogal get⁹¹st⁹²cld op zijn privacy ook. 17 Hij leeft van wat hij verdient met de sh¹⁸ch¹⁹ilderijtje²⁰j²¹es²¹ die hij maakt en verkoopt aan en paar vr²³ien²⁴den en wat verwanten²⁴ famileileden.²⁵ 23 Prutserij,²² noemt hij zijn werk²⁵ het zelf⁹⁶ 96 92 92. 99 Maar het is best wel goed. 27 Soms kan hij meer²⁹ krij³⁰en dan wat hij voor zo'n³¹ h³² kunstwerkje vraagt t³³h³⁴ beiden³⁴ 35, maar dat w³⁶il hij nooit¹²⁹ tet¹²¹ 37. 38 Hij is ook¹⁰⁸ doodeerlijk. 150 Het¹²² t¹²³ 122 is zo'n jone¹⁴¹ ge¹⁵²n waar¹⁴³ van wie¹⁵³ je alleen maar kan houden.

Figure 6: Transcription of a paragraph in session 30, displaying the sequence (numbers) of the modifications. For a legend of the colours and symbols used in this transcription, see Table 7 in Appendix A below.

²⁸ The numbers (e.g., n14) refer to the chronology of each modification made in this paragraph during this logged writing session and coincide with the @n attribute in XML (see also Appendix A).

Lastly, symbols can be added to the transcription to provide an option for colour-blind users to still be able to distinguish the different modifications (see Appendix A). Adding these symbols can also enhance readability at the borders of the insertions and deletions; for example, by distinguishing the contextual addition “verwanten” within the contextual addition “aan en [sic] paar vienden [sic] en familieleden [sic]”:

◆La▶a¹⁴³◆en▶¹⁴⁴◆ik▶¹⁴⁵◆we▶¹⁴⁶hem voor het gemak de krijkunstenaar noemen.◆¹⁴⁶
 Niemand weet hoe hij heet of waar hij vandaan komt. ◆Laat ik hem voor het gemak de krijkunstenaar
 noemen.◆¹²⁷★¹⁴◆E▶rs▶¹²⁷◆cht▶¹²⁸veel kan ik niet over hem kwijt.◆¹⁴⁸★★¹⁵Hij is hoogst
 bescheiden.★★¹⁶Nogal ge▶t▶¹²⁴s▶t▶¹²⁵eld op zijn privacy ook.★★¹⁷Hij leeft van wat hij verdient met
 de s▶h▶¹⁴⁸c▶h▶¹⁴⁹ilderijt▶e▶¹⁵⁰jes▶¹⁵¹die hij maakt en verkoopt▶aan en paar v▶r▶¹⁵²je▶n▶¹⁵³den
 en wat▶verwanten▶¹⁵⁴▶familieleden▶¹⁵⁵▶¹⁵⁶▶¹⁵⁷★★²²⁷Prutserij,' noemt hij▶zijn werk▶¹⁵⁸▶het
 zelf▶¹⁵⁹▶¹⁶⁰▶¹⁶¹▶¹⁶²▶¹⁶³▶¹⁶⁴▶¹⁶⁵▶¹⁶⁶▶¹⁶⁷▶¹⁶⁸▶¹⁶⁹▶¹⁷⁰▶¹⁷¹▶¹⁷²▶¹⁷³▶¹⁷⁴▶¹⁷⁵▶¹⁷⁶▶¹⁷⁷▶¹⁷⁸▶¹⁷⁹▶¹⁸⁰▶¹⁸¹▶¹⁸²▶¹⁸³▶¹⁸⁴▶¹⁸⁵▶¹⁸⁶▶¹⁸⁷▶¹⁸⁸▶¹⁸⁹▶¹⁹⁰▶¹⁹¹▶¹⁹²▶¹⁹³▶¹⁹⁴▶¹⁹⁵▶¹⁹⁶▶¹⁹⁷▶¹⁹⁸▶¹⁹⁹▶²⁰⁰▶²⁰¹▶²⁰²▶²⁰³▶²⁰⁴▶²⁰⁵▶²⁰⁶▶²⁰⁷▶²⁰⁸▶²⁰⁹▶²¹⁰▶²¹¹▶²¹²▶²¹³▶²¹⁴▶²¹⁵▶²¹⁶▶²¹⁷▶²¹⁸▶²¹⁹▶²²⁰▶²²¹▶²²²▶²²³▶²²⁴▶²²⁵▶²²⁶▶²²⁷▶²²⁸▶²²⁹▶²³⁰▶²³¹▶²³²▶²³³▶²³⁴▶²³⁵▶²³⁶▶²³⁷▶²³⁸▶²³⁹▶²⁴⁰▶²⁴¹▶²⁴²▶²⁴³▶²⁴⁴▶²⁴⁵▶²⁴⁶▶²⁴⁷▶²⁴⁸▶²⁴⁹▶²⁵⁰▶²⁵¹▶²⁵²▶²⁵³▶²⁵⁴▶²⁵⁵▶²⁵⁶▶²⁵⁷▶²⁵⁸▶²⁵⁹▶²⁶⁰▶²⁶¹▶²⁶²▶²⁶³▶²⁶⁴▶²⁶⁵▶²⁶⁶▶²⁶⁷▶²⁶⁸▶²⁶⁹▶²⁷⁰▶²⁷¹▶²⁷²▶²⁷³▶²⁷⁴▶²⁷⁵▶²⁷⁶▶²⁷⁷▶²⁷⁸▶²⁷⁹▶²⁸⁰▶²⁸¹▶²⁸²▶²⁸³▶²⁸⁴▶²⁸⁵▶²⁸⁶▶²⁸⁷▶²⁸⁸▶²⁸⁹▶²⁹⁰▶²⁹¹▶²⁹²▶²⁹³▶²⁹⁴▶²⁹⁵▶²⁹⁶▶²⁹⁷▶²⁹⁸▶²⁹⁹▶³⁰⁰▶³⁰¹▶³⁰²▶³⁰³▶³⁰⁴▶³⁰⁵▶³⁰⁶▶³⁰⁷▶³⁰⁸▶³⁰⁹▶³¹⁰▶³¹¹▶³¹²▶³¹³▶³¹⁴▶³¹⁵▶³¹⁶▶³¹⁷▶³¹⁸▶³¹⁹▶³²⁰▶³²¹▶³²²▶³²³▶³²⁴▶³²⁵▶³²⁶▶³²⁷▶³²⁸▶³²⁹▶³³⁰▶³³¹▶³³²▶³³³▶³³⁴▶³³⁵▶³³⁶▶³³⁷▶³³⁸▶³³⁹▶³⁴⁰▶³⁴¹▶³⁴²▶³⁴³▶³⁴⁴▶³⁴⁵▶³⁴⁶▶³⁴⁷▶³⁴⁸▶³⁴⁹▶³⁵⁰▶³⁵¹▶³⁵²▶³⁵³▶³⁵⁴▶³⁵⁵▶³⁵⁶▶³⁵⁷▶³⁵⁸▶³⁵⁹▶³⁶⁰▶³⁶¹▶³⁶²▶³⁶³▶³⁶⁴▶³⁶⁵▶³⁶⁶▶³⁶⁷▶³⁶⁸▶³⁶⁹▶³⁷⁰▶³⁷¹▶³⁷²▶³⁷³▶³⁷⁴▶³⁷⁵▶³⁷⁶▶³⁷⁷▶³⁷⁸▶³⁷⁹▶³⁸⁰▶³⁸¹▶³⁸²▶³⁸³▶³⁸⁴▶³⁸⁵▶³⁸⁶▶³⁸⁷▶³⁸⁸▶³⁸⁹▶³⁹⁰▶³⁹¹▶³⁹²▶³⁹³▶³⁹⁴▶³⁹⁵▶³⁹⁶▶³⁹⁷▶³⁹⁸▶³⁹⁹▶⁴⁰⁰▶⁴⁰¹▶⁴⁰²▶⁴⁰³▶⁴⁰⁴▶⁴⁰⁵▶⁴⁰⁶▶⁴⁰⁷▶⁴⁰⁸▶⁴⁰⁹▶⁴¹⁰▶⁴¹¹▶⁴¹²▶⁴¹³▶⁴¹⁴▶⁴¹⁵▶⁴¹⁶▶⁴¹⁷▶⁴¹⁸▶⁴¹⁹▶⁴²⁰▶⁴²¹▶⁴²²▶⁴²³▶⁴²⁴▶⁴²⁵▶⁴²⁶▶⁴²⁷▶⁴²⁸▶⁴²⁹▶⁴³⁰▶⁴³¹▶⁴³²▶⁴³³▶⁴³⁴▶⁴³⁵▶⁴³⁶▶⁴³⁷▶⁴³⁸▶⁴³⁹▶⁴⁴⁰▶⁴⁴¹▶⁴⁴²▶⁴⁴³▶⁴⁴⁴▶⁴⁴⁵▶⁴⁴⁶▶⁴⁴⁷▶⁴⁴⁸▶⁴⁴⁹▶⁴⁵⁰▶⁴⁵¹▶⁴⁵²▶⁴⁵³▶⁴⁵⁴▶⁴⁵⁵▶⁴⁵⁶▶⁴⁵⁷▶⁴⁵⁸▶⁴⁵⁹▶⁴⁶⁰▶⁴⁶¹▶⁴⁶²▶⁴⁶³▶⁴⁶⁴▶⁴⁶⁵▶⁴⁶⁶▶⁴⁶⁷▶⁴⁶⁸▶⁴⁶⁹▶⁴⁷⁰▶⁴⁷¹▶⁴⁷²▶⁴⁷³▶⁴⁷⁴▶⁴⁷⁵▶⁴⁷⁶▶⁴⁷⁷▶⁴⁷⁸▶⁴⁷⁹▶⁴⁸⁰▶⁴⁸¹▶⁴⁸²▶⁴⁸³▶⁴⁸⁴▶⁴⁸⁵▶⁴⁸⁶▶⁴⁸⁷▶⁴⁸⁸▶⁴⁸⁹▶⁴⁹⁰▶⁴⁹¹▶⁴⁹²▶⁴⁹³▶⁴⁹⁴▶⁴⁹⁵▶⁴⁹⁶▶⁴⁹⁷▶⁴⁹⁸▶⁴⁹⁹▶⁵⁰⁰▶⁵⁰¹▶⁵⁰²▶⁵⁰³▶⁵⁰⁴▶⁵⁰⁵▶⁵⁰⁶▶⁵⁰⁷▶⁵⁰⁸▶⁵⁰⁹▶⁵¹⁰▶⁵¹¹▶⁵¹²▶⁵¹³▶⁵¹⁴▶⁵¹⁵▶⁵¹⁶▶⁵¹⁷▶⁵¹⁸▶⁵¹⁹▶⁵²⁰▶⁵²¹▶⁵²²▶⁵²³▶⁵²⁴▶⁵²⁵▶⁵²⁶▶⁵²⁷▶⁵²⁸▶⁵²⁹▶⁵³⁰▶⁵³¹▶⁵³²▶⁵³³▶⁵³⁴▶⁵³⁵▶⁵³⁶▶⁵³⁷▶⁵³⁸▶⁵³⁹▶⁵⁴⁰▶⁵⁴¹▶⁵⁴²▶⁵⁴³▶⁵⁴⁴▶⁵⁴⁵▶⁵⁴⁶▶⁵⁴⁷▶⁵⁴⁸▶⁵⁴⁹▶⁵⁵⁰▶⁵⁵¹▶⁵⁵²▶⁵⁵³▶⁵⁵⁴▶⁵⁵⁵▶⁵⁵⁶▶⁵⁵⁷▶⁵⁵⁸▶⁵⁵⁹▶⁵⁶⁰▶⁵⁶¹▶⁵⁶²▶⁵⁶³▶⁵⁶⁴▶⁵⁶⁵▶⁵⁶⁶▶⁵⁶⁷▶⁵⁶⁸▶⁵⁶⁹▶⁵⁷⁰▶⁵⁷¹▶⁵⁷²▶⁵⁷³▶⁵⁷⁴▶⁵⁷⁵▶⁵⁷⁶▶⁵⁷⁷▶⁵⁷⁸▶⁵⁷⁹▶⁵⁸⁰▶⁵⁸¹▶⁵⁸²▶⁵⁸³▶⁵⁸⁴▶⁵⁸⁵▶⁵⁸⁶▶⁵⁸⁷▶⁵⁸⁸▶⁵⁸⁹▶⁵⁹⁰▶⁵⁹¹▶⁵⁹²▶⁵⁹³▶⁵⁹⁴▶⁵⁹⁵▶⁵⁹⁶▶⁵⁹⁷▶⁵⁹⁸▶⁵⁹⁹▶⁶⁰⁰▶⁶⁰¹▶⁶⁰²▶⁶⁰³▶⁶⁰⁴▶⁶⁰⁵▶⁶⁰⁶▶⁶⁰⁷▶⁶⁰⁸▶⁶⁰⁹▶⁶¹⁰▶⁶¹¹▶⁶¹²▶⁶¹³▶⁶¹⁴▶⁶¹⁵▶⁶¹⁶▶⁶¹⁷▶⁶¹⁸▶⁶¹⁹▶⁶²⁰▶⁶²¹▶⁶²²▶⁶²³▶⁶²⁴▶⁶²⁵▶⁶²⁶▶⁶²⁷▶⁶²⁸▶⁶²⁹▶⁶³⁰▶⁶³¹▶⁶³²▶⁶³³▶⁶³⁴▶⁶³⁵▶⁶³⁶▶⁶³⁷▶⁶³⁸▶⁶³⁹▶⁶⁴⁰▶⁶⁴¹▶⁶⁴²▶⁶⁴³▶⁶⁴⁴▶⁶⁴⁵▶⁶⁴⁶▶⁶⁴⁷▶⁶⁴⁸▶⁶⁴⁹▶⁶⁵⁰▶⁶⁵¹▶⁶⁵²▶⁶⁵³▶⁶⁵⁴▶⁶⁵⁵▶⁶⁵⁶▶⁶⁵⁷▶⁶⁵⁸▶⁶⁵⁹▶⁶⁶⁰▶⁶⁶¹▶⁶⁶²▶⁶⁶³▶⁶⁶⁴▶⁶⁶⁵▶⁶⁶⁶▶⁶⁶⁷▶⁶⁶⁸▶⁶⁶⁹▶⁶⁷⁰▶⁶⁷¹▶⁶⁷²▶⁶⁷³▶⁶⁷⁴▶⁶⁷⁵▶⁶⁷⁶▶⁶⁷⁷▶⁶⁷⁸▶⁶⁷⁹▶⁶⁸⁰▶⁶⁸¹▶⁶⁸²▶⁶⁸³▶⁶⁸⁴▶⁶⁸⁵▶⁶⁸⁶▶⁶⁸⁷▶⁶⁸⁸▶⁶⁸⁹▶⁶⁹⁰▶⁶⁹¹▶⁶⁹²▶⁶⁹³▶⁶⁹⁴▶⁶⁹⁵▶⁶⁹⁶▶⁶⁹⁷▶⁶⁹⁸▶⁶⁹⁹▶⁷⁰⁰▶⁷⁰¹▶⁷⁰²▶⁷⁰³▶⁷⁰⁴▶⁷⁰⁵▶⁷⁰⁶▶⁷⁰⁷▶⁷⁰⁸▶⁷⁰⁹▶⁷¹⁰▶⁷¹¹▶⁷¹²▶⁷¹³▶⁷¹⁴▶⁷¹⁵▶⁷¹⁶▶⁷¹⁷▶⁷¹⁸▶⁷¹⁹▶⁷²⁰▶⁷²¹▶⁷²²▶⁷²³▶⁷²⁴▶⁷²⁵▶⁷²⁶▶⁷²⁷▶⁷²⁸▶⁷²⁹▶⁷³⁰▶⁷³¹▶⁷³²▶⁷³³▶⁷³⁴▶⁷³⁵▶⁷³⁶▶⁷³⁷▶⁷³⁸▶⁷³⁹▶⁷⁴⁰▶⁷⁴¹▶⁷⁴²▶⁷⁴³▶⁷⁴⁴▶⁷⁴⁵▶⁷⁴⁶▶⁷⁴⁷▶⁷⁴⁸▶⁷⁴⁹▶⁷⁵⁰▶⁷⁵¹▶⁷⁵²▶⁷⁵³▶⁷⁵⁴▶⁷⁵⁵▶⁷⁵⁶▶⁷⁵⁷▶⁷⁵⁸▶⁷⁵⁹▶⁷⁶⁰▶⁷⁶¹▶⁷⁶²▶⁷⁶³▶⁷⁶⁴▶⁷⁶⁵▶⁷⁶⁶▶⁷⁶⁷▶⁷⁶⁸▶⁷⁶⁹▶⁷⁷⁰▶⁷⁷¹▶⁷⁷²▶⁷⁷³▶⁷⁷⁴▶⁷⁷⁵▶⁷⁷⁶▶⁷⁷⁷▶⁷⁷⁸▶⁷⁷⁹▶⁷⁸⁰▶⁷⁸¹▶⁷⁸²▶⁷⁸³▶⁷⁸⁴▶⁷⁸⁵▶⁷⁸⁶▶⁷⁸⁷▶⁷⁸⁸▶⁷⁸⁹▶⁷⁹⁰▶⁷⁹¹▶⁷⁹²▶⁷⁹³▶⁷⁹⁴▶⁷⁹⁵▶⁷⁹⁶▶⁷⁹⁷▶⁷⁹⁸▶⁷⁹⁹▶⁸⁰⁰▶⁸⁰¹▶⁸⁰²▶⁸⁰³▶⁸⁰⁴▶⁸⁰⁵▶⁸⁰⁶▶⁸⁰⁷▶⁸⁰⁸▶⁸⁰⁹▶⁸¹⁰▶⁸¹¹▶⁸¹²▶⁸¹³▶⁸¹⁴▶⁸¹⁵▶⁸¹⁶▶⁸¹⁷▶⁸¹⁸▶⁸¹⁹▶⁸²⁰▶⁸²¹▶⁸²²▶⁸²³▶⁸²⁴▶⁸²⁵▶⁸²⁶▶⁸²⁷▶⁸²⁸▶⁸²⁹▶⁸³⁰▶⁸³¹▶⁸³²▶⁸³³▶⁸³⁴▶⁸³⁵▶⁸³⁶▶⁸³⁷▶⁸³⁸▶⁸³⁹▶⁸⁴⁰▶⁸⁴¹▶⁸⁴²▶⁸⁴³▶⁸⁴⁴▶⁸⁴⁵▶⁸⁴⁶▶⁸⁴⁷▶⁸⁴⁸▶⁸⁴⁹▶⁸⁵⁰▶⁸⁵¹▶⁸⁵²▶⁸⁵³▶⁸⁵⁴▶⁸⁵⁵▶⁸⁵⁶▶⁸⁵⁷▶⁸⁵⁸▶⁸⁵⁹▶⁸⁶⁰▶⁸⁶¹▶⁸⁶²▶⁸⁶³▶⁸⁶⁴▶⁸⁶⁵▶⁸⁶⁶▶⁸⁶⁷▶⁸⁶⁸▶⁸⁶⁹▶⁸⁷⁰▶⁸⁷¹▶⁸⁷²▶⁸⁷³▶⁸⁷⁴▶⁸⁷⁵▶⁸⁷⁶▶⁸⁷⁷▶⁸⁷⁸▶⁸⁷⁹▶⁸⁸⁰▶⁸⁸¹▶⁸⁸²▶⁸⁸³▶⁸⁸⁴▶⁸⁸⁵▶⁸⁸⁶▶⁸⁸⁷▶⁸⁸⁸▶⁸⁸⁹▶⁸⁹⁰▶⁸⁹¹▶⁸⁹²▶⁸⁹³▶⁸⁹⁴▶⁸⁹⁵▶⁸⁹⁶▶⁸⁹⁷▶⁸⁹⁸▶⁸⁹⁹▶⁹⁰⁰▶⁹⁰¹▶⁹⁰²▶⁹⁰³▶⁹⁰⁴▶⁹⁰⁵▶⁹⁰⁶▶⁹⁰⁷▶⁹⁰⁸▶⁹⁰⁹▶⁹¹⁰▶⁹¹¹▶⁹¹²▶⁹¹³▶⁹¹⁴▶⁹¹⁵▶⁹¹⁶▶⁹¹⁷▶⁹¹⁸▶⁹¹⁹▶⁹²⁰▶⁹²¹▶⁹²²▶⁹²³▶⁹²⁴▶⁹²⁵▶⁹²⁶▶⁹²⁷▶⁹²⁸▶⁹²⁹▶⁹³⁰▶⁹³¹▶⁹³²▶⁹³³▶⁹³⁴▶⁹³⁵▶⁹³⁶▶⁹³⁷▶⁹³⁸▶⁹³⁹▶⁹⁴⁰▶⁹⁴¹▶⁹⁴²▶⁹⁴³▶⁹⁴⁴▶⁹⁴⁵▶⁹⁴⁶▶⁹⁴⁷▶⁹⁴⁸▶⁹⁴⁹▶⁹⁵⁰▶⁹⁵¹▶⁹⁵²▶⁹⁵³▶⁹⁵⁴▶⁹⁵⁵▶⁹⁵⁶▶⁹⁵⁷▶⁹⁵⁸▶⁹⁵⁹▶⁹⁶⁰▶⁹⁶¹▶⁹⁶²▶⁹⁶³▶⁹⁶⁴▶⁹⁶⁵▶⁹⁶⁶▶⁹⁶⁷▶⁹⁶⁸▶⁹⁶⁹▶⁹⁷⁰▶⁹⁷¹▶⁹⁷²▶⁹⁷³▶⁹⁷⁴▶⁹⁷⁵▶⁹⁷⁶▶⁹⁷⁷▶⁹⁷⁸▶⁹⁷⁹▶⁹⁸⁰▶⁹⁸¹▶⁹⁸²▶⁹⁸³▶⁹⁸⁴▶⁹⁸⁵▶⁹⁸⁶▶⁹⁸⁷▶⁹⁸⁸▶⁹⁸⁹▶⁹⁹⁰▶⁹⁹¹▶⁹⁹²▶⁹⁹³▶⁹⁹⁴▶⁹⁹⁵▶⁹⁹⁶▶⁹⁹⁷▶⁹⁹⁸▶⁹⁹⁹▶¹⁰⁰⁰▶¹⁰⁰¹▶¹⁰⁰²▶¹⁰⁰³▶¹⁰⁰⁴▶¹⁰⁰⁵▶¹⁰⁰⁶▶¹⁰⁰⁷▶¹⁰⁰⁸▶¹⁰⁰⁹▶¹⁰¹⁰▶¹⁰¹¹▶¹⁰¹²▶¹⁰¹³▶¹⁰¹⁴▶¹⁰¹⁵▶¹⁰¹⁶▶¹⁰¹⁷▶¹⁰¹⁸▶¹⁰¹⁹▶¹⁰²⁰▶¹⁰²¹▶¹⁰²²▶¹⁰²³▶¹⁰²⁴▶¹⁰²⁵▶¹⁰²⁶▶¹⁰²⁷▶¹⁰²⁸▶¹⁰²⁹▶¹⁰³⁰▶¹⁰³¹▶¹⁰³²▶¹⁰³³▶¹⁰³⁴▶¹⁰³⁵▶¹⁰³⁶▶¹⁰³⁷▶¹⁰³⁸▶¹⁰³⁹▶¹⁰⁴⁰▶¹⁰⁴¹▶¹⁰⁴²▶¹⁰⁴³▶¹⁰⁴⁴▶¹⁰⁴⁵▶¹⁰⁴⁶▶¹⁰⁴⁷▶¹⁰⁴⁸▶¹⁰⁴⁹▶¹⁰⁵⁰▶¹⁰⁵¹▶¹⁰⁵²▶¹⁰⁵³▶¹⁰⁵⁴▶¹⁰⁵⁵▶¹⁰⁵⁶▶¹⁰⁵⁷▶¹⁰⁵⁸▶¹⁰⁵⁹▶¹⁰⁶⁰▶¹⁰⁶¹▶¹⁰⁶²▶¹⁰⁶³▶¹⁰⁶⁴▶¹⁰⁶⁵▶¹⁰⁶⁶▶¹⁰⁶⁷▶¹⁰⁶⁸▶¹⁰⁶⁹▶¹⁰⁷⁰▶¹⁰⁷¹▶¹⁰⁷²▶¹⁰⁷³▶¹⁰⁷⁴▶¹⁰⁷⁵▶¹⁰⁷⁶▶¹⁰⁷⁷▶¹⁰⁷⁸▶¹⁰⁷⁹▶¹⁰⁸⁰▶¹⁰⁸¹▶¹⁰⁸²▶¹⁰⁸³▶¹⁰⁸⁴▶¹⁰⁸⁵▶¹⁰⁸⁶▶¹⁰⁸⁷▶¹⁰⁸⁸▶¹⁰⁸⁹▶¹⁰⁹⁰▶¹⁰⁹¹▶¹⁰⁹²▶¹⁰⁹³▶¹⁰⁹⁴▶¹⁰⁹⁵▶¹⁰⁹⁶▶¹⁰⁹⁷▶¹⁰⁹⁸▶¹⁰⁹⁹▶¹¹⁰⁰▶¹¹⁰¹▶¹¹⁰²▶¹¹⁰³▶¹¹⁰⁴▶¹¹⁰⁵▶¹¹⁰⁶▶¹¹⁰⁷▶¹¹⁰⁸▶¹¹⁰⁹▶¹¹¹⁰▶¹¹¹¹▶¹¹¹²▶¹¹¹³▶¹¹¹⁴▶¹¹¹⁵▶¹¹¹⁶▶¹¹¹⁷▶¹¹¹⁸▶¹¹¹⁹▶¹¹²⁰▶¹¹²¹▶¹¹²²▶¹¹²³▶¹¹²⁴▶¹¹²⁵▶¹¹²⁶▶¹¹²⁷▶¹¹²⁸▶¹¹²⁹▶¹¹³⁰▶¹¹³¹▶¹¹³²▶¹¹³³▶¹¹³⁴▶¹¹³⁵▶¹¹³⁶▶¹¹³⁷▶¹¹³⁸▶¹¹³⁹▶¹¹⁴⁰▶¹¹⁴¹▶¹¹⁴²▶¹¹⁴³▶¹¹⁴⁴▶¹¹⁴⁵▶¹¹⁴⁶▶¹¹⁴⁷▶¹¹⁴⁸▶¹¹⁴⁹▶¹¹⁵⁰▶¹¹⁵¹▶¹¹⁵²▶¹¹⁵³▶¹¹⁵⁴▶¹¹⁵⁵▶¹¹⁵⁶▶¹¹⁵⁷▶¹¹⁵⁸▶¹¹⁵⁹▶¹¹⁶⁰▶¹¹⁶¹▶¹¹⁶²▶¹¹⁶³▶¹¹⁶⁴▶¹¹⁶⁵▶¹¹⁶⁶▶¹¹⁶⁷▶¹¹⁶⁸▶¹¹⁶⁹▶¹¹⁷⁰▶¹¹⁷¹▶¹¹⁷²▶¹¹⁷³▶¹¹⁷⁴▶¹¹⁷⁵▶¹¹⁷⁶▶¹¹⁷⁷▶¹¹⁷⁸▶¹¹⁷⁹▶¹¹⁸⁰▶¹¹⁸¹▶¹¹⁸²▶¹¹⁸³▶¹¹⁸⁴▶¹¹⁸⁵▶¹¹⁸⁶▶¹¹⁸⁷▶¹¹⁸⁸▶¹¹⁸⁹▶¹¹⁹⁰▶¹¹⁹¹▶¹¹⁹²▶¹¹⁹³▶¹¹⁹⁴▶¹¹⁹⁵▶¹¹⁹⁶▶¹¹⁹⁷▶¹¹⁹⁸▶¹¹⁹⁹▶¹²⁰⁰▶¹²⁰¹▶¹²⁰²▶¹²⁰³▶¹²⁰⁴▶¹²⁰⁵▶¹²⁰⁶▶¹²⁰⁷▶¹²⁰⁸▶¹²⁰⁹▶¹²¹⁰▶¹²¹¹▶¹²¹²▶¹²¹³▶¹²¹⁴▶¹²¹⁵▶¹²¹⁶▶¹²¹⁷▶¹²¹⁸▶¹²¹⁹▶¹²²⁰▶¹²²¹▶¹²²²▶¹²²³▶¹²²⁴▶¹²²⁵▶¹²²⁶▶¹²²⁷▶¹²²⁸▶¹²²⁹▶¹²³⁰▶¹²³¹▶¹²³²▶¹²³³▶¹²³⁴▶¹²³⁵▶¹²³⁶▶¹²³⁷▶¹²³⁸▶¹²³⁹▶¹²⁴⁰▶¹²⁴¹▶¹²⁴²▶¹²⁴³▶¹²⁴⁴▶¹²⁴⁵▶¹²⁴⁶▶¹²⁴⁷▶¹²⁴⁸▶¹²⁴⁹▶¹²⁵⁰▶¹²⁵¹▶¹²⁵²▶¹²⁵³▶¹²⁵⁴▶¹²⁵⁵▶¹²⁵⁶▶¹²⁵⁷▶¹²⁵⁸▶¹²⁵⁹▶¹²⁶⁰▶¹²⁶¹▶¹²⁶²▶¹²⁶³▶¹²⁶⁴▶¹²⁶⁵▶¹²⁶⁶▶¹²⁶⁷▶¹²⁶⁸▶¹²⁶⁹▶¹²⁷⁰▶¹²⁷¹▶¹²⁷²▶¹²⁷³▶¹²⁷⁴▶¹²⁷⁵▶¹²⁷⁶▶¹²⁷⁷▶¹²⁷⁸▶¹²⁷⁹▶¹²⁸⁰▶¹²⁸¹▶¹²⁸²▶¹²⁸³▶¹²⁸⁴▶¹²⁸⁵▶¹²⁸⁶▶¹²⁸⁷▶¹²⁸⁸▶¹²⁸⁹▶¹²⁹⁰▶¹²⁹¹▶¹²⁹²▶¹²⁹³▶¹²⁹⁴▶¹²⁹⁵▶¹²⁹⁶▶¹²⁹⁷▶¹²⁹⁸▶¹²⁹⁹▶

the stratification of corrections, deletions and additions can give insights into an author's way of working, into the work itself, the evolution of the author's *Weltanschauung*, the meaning/interpretation of the text.

(Pierazzo 2009, 171)

Compared with analogue text genetic material, the keystroke data encompasses detailed information about the process in which the text was produced. Logging writing processes with a keystroke logger enables an analysis of the textual genesis at a finer granularity: at the level of the work's nanogenesis. Future research will have to indicate whether such a nanogenetic perspective will lead to new perspectives on (the genesis of) a text and the way present-day authors write their texts. In addition, born-digital writing processes alter our notions of "variants" and "versions" (Van Hulle 2019). The transcriptions of the keystroke logging data offer a starting point for reflection on this question and for help in redefining these key concepts. In other words, there are still challenges aplenty for textual scholars in the twenty-first century.

Appendix

A. Genetic Visualization Notation Legend

In the transcriptions that were visualized in Figures 2, 3, 4, 5, 6 and 7 (see pages 26 and 30–32 above), the different types of modifications are indicated with different colours, and, if preferred, with symbols (see Table 7 below). The numbers in superscript refer to the chronology of each modification made during the logged writing session and coincide with the *@n* attribute in XML. The numbers belong to nearest textual element in the same colour, and the numbers associated with the insertion of new text and the continuation of unfinished sentences are positioned at the beginning of the relevant text segment, while those associated with the revision types are located at its end. Since insertions can be made within insertions, higher numbers can appear within text elements which have been allocated another (lower) number. The same applies to deletions within previous inserted text. If not interrupted by a number and text segment in orange (indicating the continuation of an unfinished sentence) or text and number in red and bright green (respectively pre-contextual deletions and pre-contextual additions) the production of a new sentence was uninterrupted and therefore runs from start to end.

Symbol	Colour	Indicates
★	medium violet red	newly produced text
☆	orange	returns to the leading edge
◀	bright green	pre-contextual additions
▶	dark green	contextual additions
◀	bright red	pre-contextual deletions
▶	dark red	contextual deletions
△	blue	typo-related deletions
▲	indigo	typo-related additions
◇	light blue	relocation of text (deletion)
◆	light blue	relocation of text (addition)

Table 7: Legend

B. Inputlog’s General Analysis

Table 8 below shows the General Analysis of the writing process of following sentence in Gie Bogaert’s *Roosevelt* (referenced in Table 1, p.10; Figure 2, p.26; Example 2, p.21; and Example 3, p.23):

Soms kan hij meer krijgen dan wat hij voor zo’n kunstwerkje vraagt, maar dat wil hij nooit.

In the table, the first column (*#id*) shows the number of the event (consecutively). In the second column the *Event Type* indicates the kind of event that was recorded, be that of the type *keyboard*, *mouse*, *speech*, *focus*, *insert* or *replacement*. The next column then shows the event’s *Output*. In case of a keyboard event, this output records the typed letter. The position in the fourth column (*Pos.*) represents the “cursor position”. The fifth column (*Doc. Len.*) shows the “length of the document” expressed in characters. This differs from the character production represented in the sixth column (*CP*), which shows all the characters produced during all the writing sessions so far. The *Start Time* and *Start Clock* (columns seven and eight) show the time of the “key in” — respectively in milliseconds and in clock time — and the *End Time* and *End Clock* (columns nine and ten) of each “key up”. The *Action Time* (*Act. Time*; column eleven) represents the time between each key in and key up, the *Pause Time* (column

twelve) the time between two key ins. The location of the pause is shown in *Pause Location* (column thirteen).

As such, this table reproduces all the information from Inputlog's logging output, except for three columns: *x*, *y*, and another *Pause Location* column (of the same name). In the original output, the *x* and *y* columns respectively track the location of the mouse on the *x* and *y*-axes on the screen (Leijten and Van Waes 2013), and are only logged for mouse event types (and therefore remained empty for this writing action). The other *Pause Location* column was logged a number code for each of the (written out) location types (such as BEFORE WORDS, WITHIN WORDS, etc.) — and therefore carries no additional relevant information. Overall, then, the General Analysis provides information about what was written where and when, and therefore provides all the details needed for a fine-grained reconstruction of the writing process.

Table 8

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock	Act. Time	Pause Time	Pause Location
1342	keyboard	S	4155	6702	6737	284827	00:04:44.827	284889	00:04:44.889	265	655	BEFORE WORDS
1343	keyboard	o	4156	6703	6738	285154	00:04:45.154	285217	00:04:45.217	63	327	WITHIN WORDS
1344	keyboard	m	4157	6704	6739	285279	00:04:45.279	285357	00:04:45.357	78	125	WITHIN WORDS
1345	keyboard	s	4158	6705	6740	285310	00:04:45.310	285388	00:04:45.388	78	31	WITHIN WORDS
1346	keyboard	SPACE	4159	6706	6741	285498	00:04:45.498	285560	00:04:45.560	62	188	AFTER WORDS
1347	keyboard	b	4160	6707	6742	285685	00:04:45.685	285747	00:04:45.747	62	187	BEFORE WORDS
1348	keyboard	e	4161	6708	6743	285825	00:04:45.825	285888	00:04:45.888	63	140	WITHIN WORDS
1349	keyboard	i	4162	6709	6744	285903	00:04:45.903	285966	00:04:45.966	63	78	WITHIN WORDS
1350	keyboard	d	4163	6710	6745	286044	00:04:46.044	286122	00:04:46.122	78	141	WITHIN WORDS
1351	keyboard	e	4164	6711	6746	286231	00:04:46.231	286278	00:04:46.278	47	187	WITHIN WORDS
1352	keyboard	n	4165	6712	6747	286356	00:04:46.356	286418	00:04:46.418	62	125	WITHIN WORDS
1353	keyboard	SPACE	4166	6713	6748	286590	00:04:46.590	286636	00:04:46.636	46	234	AFTER WORDS
1368	keyboard	LEFT	4160	6714	6749	298368	00:04:58.368	298399	00:04:58.399	31	11778	UNKNOWN
1370	keyboard	RIGHT	4159	6714	6749	298773	00:04:58.773	298898	00:04:58.898	125	405	UNKNOWN
1372	keyboard	RIGHT	4160	6714	6749	298976	00:04:58.976	299085	00:04:59.085	109	203	UNKNOWN
1374	keyboard	LEFT	4161	6714	6749	299319	00:04:59.319	299429	00:04:59.429	110	343	UNKNOWN
1376	keyboard	k	4160	6714	6749	300505	00:05:00.505	300567	00:05:00.567	62	1186	BEFORE WORDS
1377	keyboard	r	4161	6715	6750	301098	00:05:01.098	301145	00:05:01.145	47	593	WITHIN WORDS
1378	keyboard	i	4162	6716	6751	301691	00:05:01.691	301769	00:05:01.769	78	593	WITHIN WORDS
1379	keyboard	j	4163	6717	6752	301800	00:05:01.800	301878	00:05:01.878	78	109	WITHIN WORDS
1380	keyboard	g	4164	6718	6753	302096	00:05:02.096	302143	00:05:02.143	47	296	WITHIN WORDS
1381	keyboard	t	4165	6719	6754	302237	00:05:02.237	302299	00:05:02.299	62	141	WITHIN WORDS
1382	keyboard	SPACE	4166	6720	6755	302455	00:05:02.455	302502	00:05:02.502	47	218	AFTER WORDS
1383	keyboard	h	4167	6721	6756	302736	00:05:02.736	302798	00:05:02.798	62	281	BEFORE WORDS
1384	keyboard	LEFT	4168	6722	6757	304483	00:05:04.483	304545	00:05:04.545	62	1747	UNKNOWN
1386	keyboard	LEFT	4167	6722	6757	304670	00:05:04.670	304717	00:05:04.717	47	187	UNKNOWN
1388	keyboard	LEFT	4166	6722	6757	304811	00:05:04.811	304857	00:05:04.857	46	141	UNKNOWN
1390	keyboard	LEFT	4165	6722	6757	304967	00:05:04.967	305013	00:05:05.013	46	156	UNKNOWN
1392	keyboard	LEFT	4164	6722	6757	305107	00:05:05.107	305169	00:05:05.169	62	140	UNKNOWN
1394	keyboard	LEFT	4163	6722	6757	305263	00:05:05.263	305325	00:05:05.325	62	156	UNKNOWN
1396	keyboard	LEFT	4162	6722	6757	305419	00:05:05.419	305497	00:05:05.497	78	156	UNKNOWN
1398	keyboard	LEFT	4161	6722	6757	305622	00:05:05.622	305684	00:05:05.684	62	203	UNKNOWN

Table 8: continued

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock	Act. Time	Pause Time	Pause Location
1400	keyboard	k	4160	6722	6757	306480	00:05:06.480	306527	00:05:06.527	47	858	BEFORE WORDS
1401	keyboard	a	4161	6723	6758	306558	00:05:06.558	306651	00:05:06.651	93	78	WITHIN WORDS
1402	keyboard	n	4162	6724	6759	306667	00:05:06.667	306729	00:05:06.729	62	109	WITHIN WORDS
1403	keyboard	SPACE	4163	6725	6760	306839	00:05:06.839	306885	00:05:06.885	46	172	AFTER WORDS
1404	keyboard	h	4164	6726	6761	307041	00:05:07.041	307088	00:05:07.088	47	202	BEFORE WORDS
1405	keyboard	i	4165	6727	6762	307197	00:05:07.197	307307	00:05:07.307	110	156	WITHIN WORDS
1406	keyboard	j	4166	6728	6763	307291	00:05:07.291	307369	00:05:07.369	78	94	WITHIN WORDS
1407	keyboard	SPACE	4167	6729	6764	307494	00:05:07.494	307572	00:05:07.572	78	203	AFTER WORDS
1408	keyboard	m	4168	6730	6765	307650	00:05:07.650	307712	00:05:07.712	62	156	BEFORE WORDS
1409	keyboard	e	4169	6731	6766	307790	00:05:07.790	307868	00:05:07.868	78	140	WITHIN WORDS
1410	keyboard	e	4170	6732	6767	307931	00:05:07.931	308009	00:05:08.009	78	141	WITHIN WORDS
1411	keyboard	r	4171	6733	6768	308009	00:05:08.009	308055	00:05:08.055	46	78	WITHIN WORDS
1412	keyboard	SPACE	4172	6734	6769	308196	00:05:08.196	308274	00:05:08.274	78	187	AFTER WORDS
1413	keyboard	RIGHT	4173	6735	6770	309849	00:05:09.849	309912	00:05:09.912	63	1653	UNKNOWN
1415	keyboard	RIGHT	4174	6735	6770	310037	00:05:10.037	310068	00:05:10.068	31	188	UNKNOWN
1417	keyboard	RIGHT	4175	6735	6770	310161	00:05:10.161	310271	00:05:10.271	110	124	UNKNOWN
1419	keyboard	RIGHT	4176	6735	6770	310349	00:05:10.349	310458	00:05:10.458	109	188	UNKNOWN
1421	keyboard	RIGHT	4177	6735	6770	310707	00:05:10.707	310785	00:05:10.785	78	358	UNKNOWN
1423	keyboard	e	4178	6735	6770	311409	00:05:11.409	311487	00:05:11.487	78	702	BEFORE WORDS
1424	keyboard	n	4179	6736	6771	311550	00:05:11.550	311597	00:05:11.597	47	141	WITHIN WORDS
1425	keyboard	SPACE	4180	6737	6772	311706	00:05:11.706	311784	00:05:11.784	78	156	AFTER WORDS
1426	keyboard	d	4181	6738	6773	311831	00:05:11.831	311909	00:05:11.909	78	125	BEFORE WORDS
1427	keyboard	a	4182	6739	6774	311987	00:05:11.987	312065	00:05:12.065	78	156	WITHIN WORDS
1428	keyboard	n	4183	6740	6775	312096	00:05:12.096	312143	00:05:12.143	47	109	WITHIN WORDS
1429	keyboard	SPACE	4184	6741	6776	312252	00:05:12.252	312330	00:05:12.330	78	156	AFTER WORDS
1430	keyboard	w	4185	6742	6777	312392	00:05:12.392	312486	00:05:12.486	94	140	BEFORE WORDS
1431	keyboard	a	4186	6743	6778	312548	00:05:12.548	312626	00:05:12.626	78	156	WITHIN WORDS
1432	keyboard	t	4187	6744	6779	312626	00:05:12.626	312689	00:05:12.689	63	78	WITHIN WORDS
1433	keyboard	SPACE	4188	6745	6780	312829	00:05:12.829	312907	00:05:12.907	78	203	AFTER WORDS
1434	keyboard	h	4189	6746	6781	313188	00:05:13.188	313250	00:05:13.250	62	359	BEFORE WORDS
1435	keyboard	i	4190	6747	6782	313359	00:05:13.359	313469	00:05:13.469	110	171	WITHIN WORDS
1436	keyboard	j	4191	6748	6783	313453	00:05:13.453	313547	00:05:13.547	94	94	WITHIN WORDS
1437	keyboard	SPACE	4192	6749	6784	313687	00:05:13.687	313765	00:05:13.765	78	234	AFTER WORDS
1438	keyboard	v	4193	6750	6785	313812	00:05:13.812	313890	00:05:13.890	78	125	BEFORE WORDS
1439	keyboard	o	4194	6751	6786	313921	00:05:13.921	313968	00:05:13.968	47	109	WITHIN WORDS
1440	keyboard	o	4195	6752	6787	314046	00:05:14.046	314108	00:05:14.108	62	125	WITHIN WORDS
1441	keyboard	r	4196	6753	6788	314155	00:05:14.155	314217	00:05:14.217	62	109	WITHIN WORDS
1442	keyboard	SPACE	4197	6754	6789	314264	00:05:14.264	314358	00:05:14.358	94	109	AFTER WORDS
1443	keyboard	z	4198	6755	6790	314451	00:05:14.451	314529	00:05:14.529	78	187	BEFORE WORDS
1444	keyboard	o	4199	6756	6791	314592	00:05:14.592	314654	00:05:14.654	62	141	WITHIN WORDS
1445	keyboard	n	4200	6757	6792	314810	00:05:14.810	314857	00:05:14.857	47	218	WITHIN WORDS
1446	keyboard	'	4201	6758	6793	314919	00:05:14.919	314966	00:05:14.966	47	109	WITHIN WORDS
1447	keyboard	SPACE	4202	6759	6794	315075	00:05:15.075	315138	00:05:15.138	63	156	AFTER WORDS
1448	keyboard	k	4203	6760	6795	315543	00:05:15.543	315606	00:05:15.606	63	468	BEFORE WORDS
1449	keyboard	u	4204	6761	6796	315746	00:05:15.746	315793	00:05:15.793	47	203	WITHIN WORDS
1450	keyboard	n	4205	6762	6797	315918	00:05:15.918	315980	00:05:15.980	62	172	WITHIN WORDS
1451	keyboard	s	4206	6763	6798	316027	00:05:16.027	316105	00:05:16.105	78	109	WITHIN WORDS
1452	keyboard	t	4207	6764	6799	316152	00:05:16.152	316183	00:05:16.183	31	125	WITHIN WORDS
1453	keyboard	w	4208	6765	6800	316277	00:05:16.277	316355	00:05:16.355	78	125	WITHIN WORDS
1454	keyboard	e	4209	6766	6801	316526	00:05:16.526	316620	00:05:16.620	94	249	WITHIN WORDS

Table 8: continued

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock	Act. Time	Pause Time	Pause Location
1455	keyboard	r	4210	6767	6802	316589	00:05:16.589	316667	00:05:16.667	78	63	WITHIN WORDS
1456	keyboard	k	4211	6768	6803	316760	00:05:16.760	316823	00:05:16.823	63	171	WITHIN WORDS
1457	keyboard	j	4212	6769	6804	316916	00:05:16.916	316963	00:05:16.963	47	156	WITHIN WORDS
1458	keyboard	e	4213	6770	6805	317025	00:05:17.025	317088	00:05:17.088	63	109	WITHIN WORDS
1459	keyboard	SPACE	4214	6771	6806	317150	00:05:17.150	317213	00:05:17.213	63	125	AFTER WORDS
1460	keyboard	v	4215	6772	6807	317369	00:05:17.369	317415	00:05:17.415	46	219	BEFORE WORDS
1461	keyboard	r	4216	6773	6808	317540	00:05:17.540	317587	00:05:17.587	47	171	WITHIN WORDS
1462	keyboard	a	4217	6774	6809	317634	00:05:17.634	317696	00:05:17.696	62	94	WITHIN WORDS
1463	keyboard	a	4218	6775	6810	317790	00:05:17.790	317868	00:05:17.868	78	156	WITHIN WORDS
1464	keyboard	g	4219	6776	6811	317961	00:05:17.961	318008	00:05:18.008	47	171	WITHIN WORDS
1465	keyboard	t	4220	6777	6812	318180	00:05:18.180	318211	00:05:18.211	31	219	WITHIN WORDS
1504	keyboard	LEFT	4202	6778	6813	321581	00:05:21.581	321596	00:05:21.596	15	3401	UNKNOWN
1506	keyboard	LEFT	4201	6778	6813	321908	00:05:21.908	321986	00:05:21.986	78	327	UNKNOWN
1508	keyboard	DELETE	4200	6777	6813	322298	00:05:22.298	322329	00:05:22.329	31	390	REVISION
1509	replacement	[4200:4201]	4200	6777	6813	322298	00:05:22.298	322329	00:05:22.329	0	0	CHANGE
1511	keyboard	RIGHT	4200	6777	6813	322517	00:05:22.517	322595	00:05:22.595	78	219	UNKNOWN
1513	keyboard	n	4201	6777	6813	323546	00:05:23.546	323609	00:05:23.609	63	1029	BEFORE WORDS
1546	keyboard	RIGHT	4218	6778	6814	324903	00:05:24.903	324935	00:05:24.935	32	1357	UNKNOWN
1548	keyboard	RIGHT	4219	6778	6814	325231	00:05:25.231	325309	00:05:25.309	78	328	UNKNOWN
1550	keyboard	RIGHT	4220	6778	6814	325481	00:05:25.481	325496	00:05:25.496	15	250	UNKNOWN
1552	keyboard	DELETE	4221	6777	6814	325949	00:05:25.949	325995	00:05:25.995	46	468	REVISION
1553	replacement	[4221:4222]	4221	6777	6814	325949	00:05:25.949	325995	00:05:25.995	0	0	CHANGE
1555	keyboard	SPACE	4221	6777	6814	326448	00:05:26.448	326541	00:05:26.541	93	499	AFTER WORDS
1556	keyboard	DELETE	4222	6777	6815	326838	00:05:26.838	326916	00:05:26.916	78	390	REVISION
1557	replacement	[4222:4223]	4222	6777	6815	326838	00:05:26.838	326916	00:05:26.916	0	0	CHANGE
1559	keyboard	DELETE	4222	6776	6815	327041	00:05:27.041	327103	00:05:27.103	62	203	REVISION
1560	replacement	[4222:4223]	4222	6776	6815	327041	00:05:27.041	327103	00:05:27.103	0	0	CHANGE
1562	keyboard	DELETE	4222	6775	6815	327181	00:05:27.181	327181	00:05:27.181	0	140	REVISION
1563	replacement	[4222:4223]	4222	6775	6815	327181	00:05:27.181	327181	00:05:27.181	0	0	CHANGE
1565	keyboard	DELETE	4222	6774	6815	327696	00:05:27.696	327696	00:05:27.696	0	515	REVISION
1566	replacement	[4222:4223]	4222	6774	6815	327696	00:05:27.696	327696	00:05:27.696	0	0	CHANGE
1568	keyboard	DELETE	4222	6773	6815	327727	00:05:27.727	327727	00:05:27.727	0	31	REVISION
1569	replacement	[4222:4223]	4222	6773	6815	327727	00:05:27.727	327727	00:05:27.727	0	0	CHANGE
1571	keyboard	DELETE	4222	6772	6815	327758	00:05:27.758	327758	00:05:27.758	0	31	REVISION
1572	replacement	[4222:4223]	4222	6772	6815	327758	00:05:27.758	327758	00:05:27.758	0	0	CHANGE
1574	keyboard	DELETE	4222	6771	6815	327789	00:05:27.789	327789	00:05:27.789	0	31	REVISION
1575	replacement	[4222:4223]	4222	6771	6815	327789	00:05:27.789	327789	00:05:27.789	0	0	CHANGE
1577	keyboard	DELETE	4222	6770	6815	327836	00:05:27.836	327836	00:05:27.836	0	47	REVISION
1578	replacement	[4222:4223]	4222	6770	6815	327836	00:05:27.836	327836	00:05:27.836	0	0	CHANGE
1580	keyboard	DELETE	4222	6769	6815	327867	00:05:27.867	327899	00:05:27.899	32	31	REVISION
1581	replacement	[4222:4223]	4222	6769	6815	327867	00:05:27.867	327899	00:05:27.899	0	0	CHANGE
1583	keyboard	RETURN	4222	6769	6815	328476	00:05:28.476	328569	00:05:28.569	93	609	BEFORE PARAGRAPHS
1584	keyboard	RETURN	4223	6770	6816	328647	00:05:28.647	328710	00:05:28.710	63	171	BEFORE WORDS
1585	keyboard	LEFT	4224	6771	6817	329147	00:05:29.147	329240	00:05:29.240	93	500	UNKNOWN
1587	keyboard	LEFT	4223	6771	6817	329334	00:05:29.334	329396	00:05:29.396	62	187	UNKNOWN
1589	keyboard	LEFT	4222	6771	6817	329490	00:05:29.490	329584	00:05:29.584	94	156	UNKNOWN
1591	keyboard	LEFT	4221	6771	6817	329693	00:05:29.693	329771	00:05:29.771	78	203	UNKNOWN
1593	keyboard	RIGHT	4220	6771	6817	330020	00:05:30.020	330130	00:05:30.130	110	327	UNKNOWN
1595	keyboard	,	4221	6771	6817	330847	00:05:30.847	330894	00:05:30.894	47	827	AFTER WORDS
1596	keyboard	SPACE	4222	6772	6818	330988	00:05:30.988	331050	00:05:31.050	62	141	AFTER WORDS

Table 8: continued

#id	Event Type	Output	Pos.	Doc. Len.	CP	Start Time	Start Clock	End Time	End Clock	Act. Time	Pause Time	Pause Location
1597	keyboard	m	4223	6773	6819	331175	00:05:31.175	331237	00:05:31.237	62	187	BEFORE WORDS
1598	keyboard	a	4224	6774	6820	331206	00:05:31.206	331284	00:05:31.284	78	31	WITHIN WORDS
1599	keyboard	a	4225	6775	6821	331362	00:05:31.362	331456	00:05:31.456	94	156	WITHIN WORDS
1600	keyboard	r	4226	6776	6822	331518	00:05:31.518	331565	00:05:31.565	47	156	WITHIN WORDS
1601	keyboard	SPACE	4227	6777	6823	331736	00:05:31.736	331799	00:05:31.799	63	218	AFTER WORDS
1602	keyboard	d	4228	6778	6824	331830	00:05:31.830	331892	00:05:31.892	62	94	BEFORE WORDS
1603	keyboard	a	4229	6779	6825	331986	00:05:31.986	332064	00:05:32.064	78	156	WITHIN WORDS
1604	keyboard	t	4230	6780	6826	332111	00:05:32.111	332142	00:05:32.142	31	125	WITHIN WORDS
1605	keyboard	SPACE	4231	6781	6827	332282	00:05:32.282	332345	00:05:32.345	63	171	AFTER WORDS
1606	keyboard	i	4232	6782	6828	332579	00:05:32.579	332626	00:05:32.626	47	297	BEFORE WORDS
1607	keyboard	l	4233	6783	6829	332688	00:05:32.688	332766	00:05:32.766	78	109	WITHIN WORDS
1608	keyboard	SPACE	4234	6784	6830	332906	00:05:32.906	332938	00:05:32.938	32	218	AFTER WORDS
1609	keyboard	h	4235	6785	6831	333203	00:05:33.203	333250	00:05:33.250	47	297	BEFORE WORDS
1610	keyboard	i	4236	6786	6832	333374	00:05:33.374	333515	00:05:33.515	141	171	WITHIN WORDS
1611	keyboard	j	4237	6787	6833	333484	00:05:33.484	333577	00:05:33.577	93	110	WITHIN WORDS
1612	keyboard	SPACE	4238	6788	6834	333702	00:05:33.702	333749	00:05:33.749	47	218	AFTER WORDS
1613	keyboard	n	4239	6789	6835	333889	00:05:33.889	333952	00:05:33.952	63	187	BEFORE WORDS
1614	keyboard	i	4240	6790	6836	333998	00:05:33.998	334076	00:05:34.076	78	109	WITHIN WORDS
1615	keyboard	e	4241	6791	6837	334123	00:05:34.123	334186	00:05:34.186	63	125	WITHIN WORDS
1616	keyboard	t	4242	6792	6838	334279	00:05:34.279	334342	00:05:34.342	63	156	WITHIN WORDS
1637	keyboard	LEFT	4233	6793	6839	336775	00:05:36.775	336791	00:05:36.791	16	2496	UNKNOWN
1639	keyboard	LEFT	4232	6793	6839	337025	00:05:37.025	337087	00:05:37.087	62	250	UNKNOWN
1641	keyboard	RIGHT	4231	6793	6839	337399	00:05:37.399	337508	00:05:37.508	109	374	UNKNOWN
1643	keyboard	w	4232	6793	6839	338351	00:05:38.351	338413	00:05:38.413	62	952	BEFORE WORDS
1644	keyboard	RIGHT	4233	6794	6840	341580	00:05:41.580	341658	00:05:41.658	78	3229	UNKNOWN
1646	keyboard	RIGHT	4234	6794	6840	341767	00:05:41.767	341767	00:05:41.767	0	187	UNKNOWN
1648	keyboard	RIGHT	4235	6794	6840	342266	00:05:42.266	342266	00:05:42.266	0	499	UNKNOWN
1650	keyboard	RIGHT	4236	6794	6840	342298	00:05:42.298	342298	00:05:42.298	0	32	UNKNOWN
1652	keyboard	RIGHT	4237	6794	6840	342344	00:05:42.344	342344	00:05:42.344	0	46	UNKNOWN
1654	keyboard	RIGHT	4238	6794	6840	342376	00:05:42.376	342376	00:05:42.376	0	32	UNKNOWN
1656	keyboard	RIGHT	4239	6794	6840	342407	00:05:42.407	342407	00:05:42.407	0	31	UNKNOWN
1658	keyboard	RIGHT	4240	6794	6840	342438	00:05:42.438	342438	00:05:42.438	0	31	UNKNOWN
1660	keyboard	RIGHT	4241	6794	6840	342485	00:05:42.485	342485	00:05:42.485	0	47	UNKNOWN
1662	keyboard	RIGHT	4242	6794	6840	342516	00:05:42.516	342547	00:05:42.547	31	31	UNKNOWN
1664	keyboard	RIGHT	4243	6794	6840	343000	00:05:43	343109	00:05:43.109	109	484	UNKNOWN
1666	keyboard	.	4244	6794	6840	344216	00:05:44.216	344279	00:05:44.279	63	1216	AFTER WORDS

Bibliography

ANDREWS, TARA L. and JORIS J. VAN ZUNDERT, 2018. “What Are You Trying to Say? The Interface as an Integral Element of Argument.” In *Digital Scholarly Editions as Interfaces, Schriften des Instituts für Dokumentologie und Editorik*, volume 12, edited by Roman Bleier, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber, and Gerlinde Schneider, Norderstedt: Books on Demand (BoD), pages 3–34.

- BEAL, PETER, 2011. *A Dictionary of English Manuscript Terminology: 1450-2000*. Oxford: Oxford University Press.
- BIASI, PIERRE-MARC DE, 1996. "What is a Literary Draft? Toward a Functional Typology of Genetic Documentation." *Yale French Studies* 89, pages 26–58. doi: <https://doi.org/10.2307/2930337>.
- BLEEKER, ELLI, 2015. "The Future of the Digital Scholarly Editor. Interpretation, Subjectivity and Presence?" *Manuscritica. Revista de crítica genética* 28, pages 112–22.
- BLEIER, ROMAN and HELMUT W. KLUG, 2018. "Discussing Interfaces in Digital Scholarly Editing." In *Digital Scholarly Editions as Interfaces, Schriften des Instituts für Dokumentologie und Editorik*, volume 12, edited by Roman Bleier, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber, and Gerlinde Schneider, Norderstedt: Books on Demand (BoD), pages V–XV.
- BOGAERT, GIE, January 2013. "Gie Bogaert over zijn bijzondere manier van schrijven." Available from: <https://youtu.be/EhRFiw-RZ0Y> (Accessed: 2020–07–09).
- , 2016. *Roosevelt*. Amsterdam: De Bezige Bij.
- BURNARD, LOU, 2014. *What is the Text Encoding Initiative? How to Add Intelligent Markup to Digital Resources. New Edition* [online]. Marseille: OpenEdition Press. doi: <https://doi.org/10.4000/books.oep.426>.
- BUSCHENHENKE, FLOOR, 2016. "Het literaire werk anno 2016. Digitale schrijfprocessen vastleggen en analyseren." *Vooy's* 34(4), pages 8–20.
- CONIJN, RIANNE, MENNO VAN ZAAENEN, MARIËLLE LEIJTEN, and LUUK VAN WAES, 2019. "How to Typo? Building a Process-Based Model of Typographic Error Revisions." *The Journal of Writing Analytics* 3, pages 69–95.
- CROMBEZ, THOMAS and EDITH CASSIERS, 2017. "Postdramatic methods of adaptation in the age of digital collaborative writing." *Digital Scholarship in the Humanities* 23(1), pages 17–35. doi: <https://doi.org/10.1093/llc/fqv054>.
- DAHLSTRÖM, MATS, 2000. "Drowning by Versions." *Human IT* 4(4). Available from: <https://humanit.hb.se/article/view/174/187> (Accessed: 2021–06–30).
- DILLEN, WOUT, 2015. *Digital Scholarly Editing for the Genetic Orientation: The Making of a Genetic Edition of Samuel Beckett's Works*. DPhil, University of Antwerp, Antwerpen.

- , 2018. "The Editor in the Interface: Guiding the User through Texts and Images." In *Digital Scholarly Editions as Interfaces*, *Schriften des Instituts für Dokumentologie und Editorik*, volume 12, edited by Roman Bleier, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber, and Gerlinde Schneider, Norderstedt: Books on Demand (BoD), pages 35–59.
- FITZGERALD, J., 1987. "Research on revision in writing." *Review of Educational Research* 57(4), pages 481–506. doi: <https://doi.org/10.2307/1170433>.
- GOODING, PAUL, JOS SMITH, and JUSTINE MANN, 2019. "The forensic imagination: interdisciplinary approaches to tracing creativity in writers' born-digital archives." *Archives and Manuscripts* 47(3), pages 374–90. doi: <https://doi.org/10.1080/01576895.2019.1608837>.
- HAYLES, N. KATHERINE, 2003. "Translating Media: Why We Should Rethink Textuality." *Yale Journal of Criticism* 16(2), pages 263–90. doi: <https://doi.org/10.1353/yale.2003.0018>.
- KIRSCHENBAUM, MATTHEW G., 2008. *Mechanisms: New Media and the Forensic Imagination*. Cambridge (MA): MIT University Press.
- , 2016. *Track Changes: A Literary History of Word Processing*. Cambridge (MA): Harvard University Press.
- KIRSCHENBAUM, MATTHEW G., RICHARD OVENDEN, and GABRIELA REDWINE, 2010. "Digital Forensics and Born-Digital Content in Cultural Heritage Collections." Technical report, Council on Library and Information Resources, Arlington (TX).
- KIRSCHENBAUM, MATTHEW G. and DOUG RESIDE, 2013. "Tracking the Changes: Textual Scholarship and the Challenge of the Born Digital." In *The Cambridge Companion to Textual Scholarship*, edited by Neil Fraistat and Julia Flanders, Cambridge (MA): Cambridge University Press, pages 257–88.
- KOLLBERG, PY, 1998. *S-notation – A computer based method for studying and representing text composition*. Lic. thesis, University of Stockholm, Stockholm.
- KOLLBERG, PY and KERSTIN S. EKLUNDH, 2002. "Studying Writers' Revising Patterns with S-Notation Analysis." In *Contemporary Tools and Techniques for Studying Writing*, *Studies in Writing*, volume 10, edited by Thierry Olive, C. Michael Levy, and Gert Rijlaarsdam, Dordrecht: Springer, pages 89–104.
- LEIJTEN, MARIËLLE and LUUK VAN WAES, 2013. "Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes." *Written Communication* 30(3), pages 358–92.

- LINDGREN, EVA, YVONNE KNOPSE, and KIRK P.H. SULLIVAN, 2019a. "Researching Writing with Observational Logging Tools from 2006 to the Present." In *Observing Writing. Insights from Keystroke Logging and Handwriting, Studies in Writing*, volume 38, edited by Eva Lindgren, Kirk P.H. Sullivan, Raquel Fidalgo, and Thierry Olive, Leiden: Brill, pages 1–29.
- LINDGREN, EVA and KIRK P.H. SULLIVAN, 2006a. "Analysing Online Revision." In *Computer Keystroke Logging and Writing: Methods and Applications, Studies in Writing*, volume 18, edited by Eva Lindgren and Kirk P.H. Sullivan, Oxford: Elsevier, pages 157–88.
- , 2006b. "Writing and the Analysis of Revision." In *Computer Keystroke Logging and Writing: Methods and Applications, Studies in Writing*, volume 18, edited by Eva Lindgren and Kirk P.H. Sullivan, Oxford: Elsevier, pages 31–44.
- LINDGREN, EVA, ASBJØRG WESTUM, HANNA OUTAKOSKI, and KIRK P.H. SULLIVAN, 2019b. "Revising at the Leading Edge: Shaping Ideas or Clearing up Noise." In *Observing Writing. Insights from Keystroke Logging and Handwriting, Studies in Writing*, volume 38, edited by Eva Lindgren, Kirk P.H. Sullivan, Raquel Fidalgo, and Thierry Olive, Leiden: Brill, pages 346–65.
- MARA, MIRIAM O'KANE, 2013. "Nuala O'Faolain: new departures in textual and genetic criticism." *Irish Studies Review* 21(3), pages 342–52.
doi: <https://doi.org/10.1080/09670882.2013.808873>.
- MATHIJSEN, MARITA, 2009. "Genetic textual editing: the end of an era." In *Was ist Textkritik? zur Geschichte und Relevanz eines Zentralbegriffs der Editionswissenschaft, Beihefte zu Editio*, volume Bd. 28, edited by Gertraud Mitterauer, Werner Maria Bauer, and Sabine Hofer, Tübingen: Niemeyer.
- MILLER, KRISTVAN S and KIRK P.H. SULLIVAN, 2006. "Keystroke Logging: An Introduction." In *Computer Keystroke Logging and Writing: Methods and Applications, Studies in Writing*, volume 18, edited by Eva Lindgren and Kirk P.H. Sullivan, Oxford: Elsevier, pages 1–9.
- PIERAZZO, ELENA, 2009. "Digital Genetic Editions: The Encoding of Time in Manuscript Transcription." In *Text Editing, Print and the Digital World*, edited by Marilyn Deegan and Kathryn Sutherland, Farnham: Ashgate, pages 169–86.
- PIERAZZO, ELENA and PETER A. STOKES, 2011. "Putting the Text back into Context: A Codicological Approach to Manuscript Transcription." In *Codicology and Palaeography in the Digital Age 2, Schriften des Instituts für Dokumentologie und Editorik*, volume 3, edited by Franz Fischer, Christiane Fritze, and Georg Vogeler, Norderstedt: Books on Demand (BoD), pages 397–429.
- RIES, THORSTEN, 2017. "Philology and the digital writing process." *Genrehybriditeit in de literatuur* 9, pages 129–58.

- , 2018. "The rationale of the born-digital dossier génétique: digital forensics and the writing process with examples from the Thomas Kling Archive." *Digital Scholarship in the Humanities* 33(2), pages 424–391. doi: <https://doi.org/10.1093/llc/fqx049>.
- SCHÄUBLE, JOSHUA and HANS WALTER GABLER, 2018. "Encodings and Visualisations of Text Processes across Document Borders." In *Digital Scholarly Editions as Interfaces*, *Schriften des Instituts für Dokumentologie und Editorik*, volume 12, edited by Roman Bleier, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber, and Gerlinde Schneider, Norderstedt: Books on Demand (BoD), pages 165–91.
- STEVENSON, MARIE, ROB SCHOONEN, and KEES DE GLOPPER, 2006. "Revising in two languages: a multi-dimensional comparison of online writing revisions in L1 and FL." *Journal of Second Language Writing* 15(3), pages 201–33. doi: <https://doi.org/10.1016/j.jslw.2006.06.002>.
- SULLIVAN, HANNAH, 2013. *The Work of Revision*. Cambridge (MA): Harvard University Press.
- TAYLOR, C.M., September 2018. "C M Taylor on 'keystroke logging project' with British Library." Available from: <https://blogs.bl.uk/english-and-drama/2018/11/c-m-taylor-on-keystroke-logging-project-with-british-library.html> (Accessed: 2020-06-24).
- TEI CONSORTIUM, 2020. "TEI P5: Guidelines for Electronic Text Encoding and Interchange." Available from: <https://tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf> (Accessed: 2020-06-24).
- VAN HOORENBEECK, ERIC, TOM PAUWAERT, LUUK VAN WAES, and MARIËLLE LEIJTEN, 2015. "A generic XML-structure for logging human computer interaction." White Paper. Available from: https://www.inputlog.net/wp-content/uploads/Generic_XML_structure_version-1_3.pdf (Accessed: 2021-06-30).
- VAN HULLE, DIRK, 2014. *Modern Manuscripts: The Extended Mind and Creative Undoing from Darwin to Beckett and Beyond*. London: Bloomsbury.
- , 2016. "Modelling a Digital Scholarly Edition for Genetic Criticism: A Rapprochement." *Variants* 12-13, pages 34–56.
- , 2019. "De logica van de tekstversie in digitaal geschreven literatuur." *Tijdschrift voor Nederlandse taal en letterkunde/Journal of Dutch Linguistics and Literature* 135(4), pages 465–77.

- , 2021. "Dynamic Facsimiles: Note on the Transcription of Born-Digital Works for Genetic Criticism." *Variants* 15–16, pages 233–43.
- VAN WAES, LUUK, MARIËLLE LEIJTEN, ÅSA WENGELIN, and EVA LINDGREN, 2011. "Logging Tools to Study Digital Writing Processes." In *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*, edited by Virginia Wise Berninger, New York (NY): Psychology Press, pages 507–34.
- VÁSÁRI, MELINDA, 2019. "Securing the Literary Evidence. Some Perspectives on Digital Forensics." In *Philology in the Making. Analog/Digital Cultures of Scholarly Writing and Reading*, edited by Pál Kelemen and Nicolas Pethes, Bielefeld: transcript Verlag, pages 287–309.
- VAUTHIER, BÉNÉDICTE, 2016. "Genetic Criticism Put to the Test by Digital Technology: Sounding out the (mainly) Digital Genetic File of El Dorado by Robert Juan-Cantavella." *Variants* 12–13, pages 163–86.
- WENGELIN, ÅSA, MARK TORRANCE, KENNETH HOLMQVIST, SOL SIMPSON, DAVID GALBRAITH, VICTORIA JOHANSSON, and ROGER JOHANSSON, 2009. "Combined eye-tracking and keystroke-logging methods for studying cognitive processes in text production." *Behavior Research Methods* 41(2), pages 337–51. doi: <https://doi.org/10.3758/BRM.41.2.337>.
- WORKGROUP ON GENETIC EDITIONS, 2010. "An Encoding Model for Genetic Editions." Available from: <https://tei-c.org/Vault/TC/tcw19.html> (Accessed: 2020–06–24).