This item is the archived peer-reviewed author-version of:

Digital library history: the virtual bookcases of James Joyce and Samuel Beckett

Reference:
Full text (Publisher's DOI): https://doi.org/10.1163/15700690-12341348
To cite this reference: http://hdl.handle.net/10067/1412220151162165141
Digital Library History

The Virtual Bookcases of James Joyce and Samuel Beckett

Dirk Van Hulle

University of Antwerp, Belgium

dirk.vanhulle@uantwerpen.be

Abstract

This article investigates to what extent it is possible, not only to digitize a writer’s extant library, but also to reconstruct her or his virtual library. The central question is whether digitizing a writer’s library is not in fact a form of digitizing a library history, and whether this implies the possibility of interdisciplinary collaboration between library history and genetic criticism.

Keywords

genetic criticism – library history – Samuel Beckett – James Joyce

The research object of genetic criticism, as defined by Daniel Ferrer, is not in the first instance a material object. But of course this research is only possible if the writing process has left material traces. The genesis of a literary work is often divided into ‘exogenesis’ (the use of external source texts) and ‘endogenesis’ (the writing of the drafts properly speaking). When, in 1979, Raymonde Debray Genette coined these terms, she already pointed out that it is hard to clearly distinguish between the two categories of exo- and endogenesis, for as soon as an author makes an excerpt or a reading note, it may already have an impact on the endogenetic writing process.

1 The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-13) / ERC grant agreement no. 313609.
This phenomenon has impressively been treated in print in the Marburg edition of Georg Büchner’s works, coordinated by the general editor, Burghard Dedner. With its ingenious system of including references to historical source texts (the so-called ‘quellenbezogener Text’), it gives shape to, and at the same time problematizes, one of the most influential concepts in recent German editorial theory, Hans Zeller’s basic tenet to clearly separate ‘Befund’ and ‘Deutung’, suggesting that a strict distinction should be made between what is the ‘record’ (the documents as objective facts) and what is ‘interpretation’. Actually, Hans Zeller himself pointed out that the definition of ‘Befund’ is quite limited: ‘A separation of record and interpretation, making verification possible, is in the case even of such plain manuscripts, which are yet difficult to decipher, attainable only by means of photomechanical reproduction (record) and transcription (interpretation) of the manuscript.’ If this applies to endogenesis, it is even more difficult to draw the line between ‘Befund’ and ‘Deutung’ in the case of exogenesis.

From an editorial perspective, the Marburg edition is an excellent example of textual scholarship’s recent efforts to include as much of the exogenesis as possible or as the medium of publication (a printed edition) will allow. Its notion of a ‘Quellenbezogener Text’ (source-related text) indicates how intense the interaction can be between exo- and endogenesis.

From the viewpoint of genetic criticism, the enhanced interest in writers’ libraries – which took shape at the beginning of the twenty-first century in such collections of essays as Bibliothèques d’écrivains (2001) or Reading Notes (2004) – increasingly manifests itself in the digital medium, for instance in the digital library of...
Charles Darwin or Melville’s *Marginalia Online*, a digital catalogue of all the books owned or borrowed by Herman Melville, and a digital edition of some of the surviving copies. By means of the following cases describing the ongoing work on the digital libraries of James Joyce and Samuel Beckett, the present article examines how various degrees of uncertainty can be integrated, to what extent digitizing a writer’s library is a form of library history, and whether this implies the possibility of interdisciplinary collaboration between library history and genetic criticism.

**James Joyce’s Personal Library**

The interaction between exo- and endogenesis can be very intense, as in the case of James Joyce. At the Centre for Manuscript Genetics (University of Antwerp) we are working on a James Joyce Digital Library.

On one hand, this consists of the *extant* library, with the indication of the marginal reading traces (the remains of Joyce’s ‘extant’ library are preserved in holding institutions in Buffalo and Austin, Texas). Joyce was not a very active ‘marginalist’, but he did produce numerous reading notes in separate notebooks. In order to write his last book, *Finnegans Wake*, Joyce needed seventeen years and he filled fifty notebooks with jottings and reading notes. Many of these notebooks have been transcribed and annotated by Vincent Deane, Daniel Ferrer and Geert Lernout. The system of note-taking typically takes the following shape: Joyce read something or something was being read to him (for instance after one of his several eye operations); he took a note; very often he was eventually unable to use this note, but sometimes he managed to use it in one of his multiple drafts and after several versions, often including a pre-book publication, it ended up in the published text of *Finnegans Wake*. When Joyce wrote his drafts, he made extensive use of these notes.

---


On the other hand, it is necessary to reconstruct a virtual library. In most cases the original book from which Joyce made an excerpt is no longer part of the extant library. But thanks to Joyce’s reading notes, we are able to retrace the source text, as in the case of *Les Monstres humains* by ‘Docteur Caufeynon’ (pseudonym of Jean Fauconney).

![Screenshot of *Les Monstres Humains* in the James Joyce Digital Library.](image)

This virtual library also contains many newspaper articles. For instance, to write the drafts for chapter 3 of *Finnegans Wake*, Joyce made use of seven notebooks. Sometimes a passage of *Finnegans Wake* is almost entirely based on notes and excerpts. A most notable example is something Joyce had read in a newspaper, *The Daily Sketch*, about the so-called Bywaters case: a woman had cheated on her husband with a man called Bywaters; she encouraged Bywaters to kill her husband,

---


12 The James Joyce Digital Library is part of a TOP BOF project, funded by the University of Antwerp, developed by Tom De Keyser, Geert Lernout and Dirk Van Hulle (PI).

which he did. Bywaters was arrested and sentenced to death. Just before he was to be hanged, the newspaper launched a petition for reprieve of Bywaters and published several reactions by people who thought Bywaters should not be hanged – among them three soldiers, ‘walking together in Fleetstreet’: ‘one gave an opinion in which all concurred. It was the woman who was to blame. [...] He proved himself a man afterwards’. 14 Joyce excerpted the whole fragment (a source discovered by Vincent Deane) and he did not do anything with the excerpt for a year, until he was writing his draft for chapter 3 and thought he could use it. The source text is not preserved among Joyce’s papers.

Another example is Joyce’s frequent use of the Encyclopaedia Britannica (11th edition). To give shape to ‘Haveth Childers Everywhere’, a section of his last book Finnegans Wake, he took notes on numerous cities, gathered from entries such as ‘London’ or ‘Stockholm’ in the Encyclopaedia. 15 Unfortunately, Joyce’s copy of the Encyclopaedia Britannica has not been preserved. So, in most cases we have to reconstruct a ‘virtual’ library on the basis of his notes: a library of books that do not exist anymore, but which we know Joyce must have read at some point.

Samuel Beckett’s Personal Library
In the case of Samuel Beckett, the extant library is slightly more extensive. Mark Nixon and I received permission from the Beckett Estate to work in Beckett’s apartment in Paris for ten days and make scans of the marginalia in his books. Apart from the analysis of this library, 16 we also digitized the books, including scans of the covers, title pages and marginalia of the books that are still present in Beckett’s apartment in Paris, as well as at the University of Reading and in a few private collections. The digital library is part of the Beckett Digital Manuscript Project (BDMP), 17 which tries to trace all the manuscripts of Samuel Beckett (scattered around the world) and bring them together in a digital way.

As in Joyce’s case, Beckett sometimes made use of the *Encyclopaedia Britannica* (11th edition). Beckett’s copy is still extant among the books in his library. He used it for instance to make notes on ‘Manichaeism’ when he was directing his play *Krapp’s Last Tape*. In his theatrical notebook, held at the Special Collections of the University of Reading, he noted – for instance – that the Manichaeans’ doctrine of ‘abstinence’ was based on a strict separation of light and darkness, each of which had its own ‘emblems’. Beckett carefully listed the five ‘Light emblems’ and the five ‘Darkness emblems’, derived from this sentence in the *Encyclopaedia* entry on ‘Manichaeism’: ‘As the earth of light has five tokens (the mild zephyr, cooling wind, bright light, quickening fire, and clear water), so has the earth of darkness also five (mist, heat, the sirocco, darkness and vapour).’ The relevant page in Beckett’s copy of the *Encyclopaedia Britannica* is marked by means of a large dog-ear.

![Dog-earred page in Beckett’s copy of the *Encyclopaedia Britannica* (11th edition).](image.png)
Darwin’s *On the Origin of Species*, which also contains dog-ears of the same size. These reading traces are recorded in the Beckett Digital Library, as part of the Beckett Digital Manuscript Project (BDMP).

Fig. 3: Dog-eared page in Beckett’s copy of Charles Darwin’s *On the Origin of Species*.

As indicated above, the extant library of Samuel Beckett may be slightly more extensive than Joyce’s personal library, but ‘The danger is in the neatness of identifications’, as Beckett wrote in his first published essay, ‘Dante…Bruno. Vico…Joyce’. In the case of Beckett’s works, the danger of this digital integration of the extant library in an edition is that it may create the impression that the exogenetic dimension of Beckett’s works is limited to the volumes presented in the digitized library. Evidently, the extent of Beckett’s reading reaches far beyond the c.700 books preserved in his apartment. Beckett gave many books to his friends or he borrowed books from libraries, especially in the 1930s. The extant library therefore shows numerous lacunae. Nevertheless, no matter how conscious we are of the lacunae, this awareness and the resulting caution are no reasons to withhold the information we find in the margins of the books we had the opportunity to get access to.

This digital library of Samuel Beckett also includes a virtual library next to the extant library: Beckett’s student library contains scans of the books he read and used while he studied and lectured at Trinity College, Dublin. To construct this virtual student library, it was necessary to consult Beckett’s letters, reading notes and lists of required reading, all of which belong to what S.E. Gontarski has called the ‘grey canon’ (the textual materials surrounding the texts that constitute the canon of Beckett’s published works). Since the reconstruction of such a library is often a matter of interpretation, a scale is useful to indicate the degree of uncertainty. For instance, about some works all the details of publication are known, even though the actual copy is no longer extant; about other works we are less certain as to the exact edition Beckett had access to.

To illustrate the useful exchange of ideas between textual scholarship and genetic criticism in this matter, the categorization of reading traces can be modeled after the typology of medieval manuscripts, proposed by Olivier Delsaux. Starting from the question ‘What is an autograph?’ Delsaux examined the relatively rare phenomenon of an autograph manuscript among medieval manuscripts that are still preserved, proceeding by means of ‘yes’ or ‘no’ questions:

<table>
<thead>
<tr>
<th>Was the manuscript authorized by the author (author’s authorization)?</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>If so, was it supervised by the author (author’s supervision)?</td>
<td>Yes</td>
</tr>
<tr>
<td>If so, was there any intervention by the author (author’s intervention)?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- (1) If it was written only partially by the author it is called an ‘autograph manufacture’ (corrected by the author)
- (2) If it was written entirely by the author, it is called an ‘autograph

---

20 Olivier Delsaux (Université catholique de Louvain-la-neuve) presented his typological terminology at a LECTIO colloquium ‘Seeing the Author at Work: Challenges of Editing an Autograph’ (Leuven, 24 March 2014). See O. Delsaux, Manuscrits et pratiques autographes chez les écrivains français de la fin du Moyen Âge. L’exemple de Christine de Pizan (Genève 2013).
21 If the manuscript is not supervised, only authorized by the author, Delsaux suggests the term ‘original manuscript’.
22 If the manuscript is only supervised by the author (without authorial intervention), Delsaux suggests the term ‘auctorial manuscript’.
manuscript’ (written or copied by the author)

Similarly, one can devise a rough typology of the kinds of reading traces one may 
encounter and take into account in a digital reconstruction of a writer’s (extant + 
virtual) library.

**Degrees of Uncertainty**

It seems valuable to create a virtual library, but it is equally important not to create the 
impression that the history of every single item in the library can be reconstructed 
with the same degree of certainty. Hence this set of criteria to mark the degree of 
uncertainty, roughly based on three categories:

1. In some cases, a book is still extant in the library without marginalia, but we 
   found reading notes in a separate notebook. This often means that Beckett first 
   read the book in a borrowed copy, when he did not have enough money to buy his 
   own books. He jotted down a few notes. Later on, he purchased the book when he 
   had the money to purchase it (usually only from the mid-1950s onwards, after the 
   success of *En attendant Godot*). This scenario applies to his 1958 Pléiade edition 
   of Montaigne’s *Essais*.  

23 Beckett jotted down one of his favourite lines from Petrarch in his so-called ‘Sam Francis’ Notebook 
(around 1950). The line was excerpted from the essay ‘De la tristesse’ by Montaigne, but before the 
acquisition of the 1958 Pléiade edition of Montaigne’s *Essais*.  

(2) Other useful sources to reconstruct the virtual library are letters and interviews. If 
a letter mentions a book that is still extant with marginalia and on which the 
author has taken notes in a separate notebook, the letter can be useful in dating the 
reading, for instance Beckett’s reading of Darwin’s *On the Origin of Species*. In 
Beckett’s case, the ‘calendars’ or lists of required reading, issued every year by 
Trinity College Dublin, constituted a useful source to date some of the books in 
the extant library and, above all, to reconstruct the virtual ‘student library’.  

24 For instance, at Trinity College, Beckett studied French literature and had to read 
several books by Honoré de Balzac. Most of them are no longer preserved in the 

extant library. In the Beckett Digital Manuscript Project, they are included in the 

---

23 Beckett jotted down one of his favourite lines from Petrarch in his so-called ‘Sam Francis’ Notebook 
(around 1950). The line was excerpted from the essay ‘De la tristesse’ by Montaigne, but before the 
acquisition of the 1958 Pléiade edition of Montaigne’s *Essais*.  

24 See also Veronica Bãlã’s doctoral dissertation, *Samuel Beckett’s Student Library and the Modern 
Novel*, defended at the University of Antwerp on 12 December 2014, supervisor D. Van Hulle 
(University of Antwerp, 2014).
‘virtual’ student library.

(3) Literary allusions in the published works can also serve as suggestions in order to reconstruct the virtual library. For instance, in the novel *Malone meurt*, there is a family that is called ‘les Louis’ and in Beckett’s own English translation of his novel, *Malone Dies*, this family is called ‘the Lamberts’. Beckett seems to be alluding to the title of one of Balzac’s novels, *Louis Lambert*. There is no copy of this novel in the extant library, but the title is mentioned among the lists of required reading at Trinity College Dublin. So, the combination of evidence enhances the degree of certainty. But if an allusion in the published novel is the only category we can rely on, the degree of certainty is much lower, and we should indicate the caution with which this item should be treated. An example would be Balzac’s *Le Colonel Chabert* (a novel about a man who goes to war; is recorded as dead; he turns out not to be dead after all; returns home; but his wife has married another man) which seems to be alluded to in Beckett’s parody of the story of the returning soldier toward the end of *L’Innommable / The Unnamable*:

They love each other, marry, in order to love each other better, more conveniently, he goes to the wars, he dies at the wars, she weeps, with emotion, at having loved him, at having lost him, yep, marries again, in order to love again, more conveniently again, they love each other, you love as many times as necessary, as necessary in order to be happy, he comes back, the other comes back, from the wars, he didn’t die at the wars after all [...] there’s a story for you [...] .

The reference to the other novel by Balzac (*Louis Lambert*) in Beckett’s previous novel, *Malone Dies*, makes the allusion to Balzac’s *Le Colonel Chabert* plausible, but it will depend on the editors of the digital library whether this type of work can be included or whether such an inclusion would imply too much interpretation (‘Deutung’). This decision is related to the discussion about the difference between what (in francophone criticism) used to be called ‘critique des sources’ on one hand and ‘intertextuality’ on the other hand. Whereas the former focuses on the detection of

---

an author’s active and conscious insertion of (literary and other) allusions in her or his text, the notion of intertextuality shifts the emphasis to the role of the reader in the creation of meaning, including the process of finding interconnections between texts.26 Readers should be enabled to fully explore a work’s intertextuality, but an editor’s own intertextual exploration should not be imposed upon them. Hence the importance of marking the degree of uncertainty.

The question then is: how do we mark this degree of uncertainty in a digital genetic edition. A possibility would be to add a “cert” attribute to the “note” tag in the mark up.27 The Beckett Digital Manuscript Project works with annotations (‘notes’ in the XML-TEI encoding of the transcriptions). For instance, at the back of Beckett’s manuscript of the novel L’Innommable, there is a reference to Francis Bacon (and indirectly to Immanuel Kant’s Werke (vol. 11). In the XML encoding, the note reads as follows:

```xml
<seg xml:id="msl_id025460" type="peralipomena" n="MS-HNC-SB-3-10,{1016}" version="1" zone="zonem3097"> De nobis ipsis illeemus { }
  <unclear>
    Bacon, Intro.
    <lb rend="lt"/>
    Novum Organon
  </unclear>
</seg>
<note type="intertextual">
  Quotation from Francis Bacon’s Novum Organon, taken as a motto by Immanuel Kant. In Beckett’s copy of Ernst Cassirer’s introduction to Kant’s Werke (which arrived in Paris in early January 1938; LER I 581), Beckett marked the sentence ‘Das Wort ‘De nobis ipsis illeemus’, das er aus Bacon entnimmt, um es der ‘Kritik der reinen Vernunft’ als Motto voranzusetzen, tritt nun mehr und mehr in Kraft’ (Cassirer in Kant 1921, XII.5). The quotation is incorporated in the text of the manuscript on MS-HNC-SB-3-10,
</note>
</seg>
</p>
```

26 After the introduction of the notion of intertextuality by such critics as Gérard Genette and Julia Kristéva, the notion became so widespread that other critics have expressed their concern about the use of the term, which tends to eclipse the role of the author in specific instances of ‘allusion’ and ‘parody’. See L. Hutcheon, A Theory of Parody. The Teachings of Twentieth-Century Art Forms (New York 1985); W. Irwin, ‘Against Intertextuality’, Philosophy and Literature, 28 (2004), pp. 227–42.

27 The Beckett Digital Manuscript Project works with Extensible Markup Language (XML) for the transcriptions of the manuscripts and follows the guidelines of the Text Encoding Initiative (TEI). Primary materials such as modern manuscripts are usually transcribed using a combination of two international standards: XML and TEI. The eXtensible Markup Language (XML) is a plain text file format for structured data. All text is surrounded by ‘tags’ that record information on the structure (chapters, paragraphs, sentences) and the features (underlined, added, deleted words) of the text. The tagset used is P5 (proposal 5) of the Text Encoding Initiative (TEI), with a few customizations for the Beckett Digital Manuscript Project. The TEI is widely used in Digital Scholarly Editions. The combination of the open file format XML and the TEI international standard for digitization guarantees the sustainability and reusability of the data.
Fig. 4: Fragment from the XML-TEI encoding of a note in the Beckett Digital Manuscript Project.

The digital infrastructure of the Beckett Digital Manuscript Project enables readers to link the manuscript to the relevant page in Beckett’s marked copy of Kant’s *Werke* (vol. 11):

Fig. 5: Marked page in Beckett’s copy of Immanuel Kant’s *Werke* (vol. 11), relating to the Latin quotation ‘De nobis ipsis silemus’ in the manuscript of *L’Innommable*.

Among the search options, the BDMP also offers a few ‘suggested searches’ (for instance, users interested in doodles can search for doodles of a particular shape). One of these suggested searches is the option ‘allusions’. But we restrict ourselves to the ones we are really certain about, that is, the ones we can corroborate with material traces. In concrete terms, this implies for instance that this option does not include the potential reference to Balzac’s *Le Colonel Chabert*.

In the XML, these allusions are marked up with a “note” tag, and a “type” attribute “intertextual” (<note type="intertextual">). To this
encoding we can add a "cert" attribute, with various possible values (high; medium; low; unknown28): <note type="intertextual" cert="high">

In order to visualize this mark up, it is possible – on one hand – to present the intertextual references with high / medium / low certainty in a particular order, or to present only the references with ‘high’ certainty. On the other hand, editors need to notify readers that these references may not be all there is, for instance by adding a disclaimer that the list of allusions is not exhaustive, and by inviting readers to contribute to the project by suggesting other intertextual readings. In this way the digital genetic edition draws readers’ attention to the possibility of editorial ‘blind spots’. This form of enhancing textual awareness seems just as crucial as the option of highlighting as many references as possible.

Probably the proposed criteria will need to be refined when tested on other writers’ libraries, but they may already offer some help with regard to the inclusion of writers’ libraries in digital genetic editions, marking the degree of uncertainty. From the perspective of library history, the question is whether digitizing a writer’s library is not in fact a form of digitizing a library history, and whether this implies the possibility of interdisciplinary collaboration between library history and genetic criticism. After all, most of the time we are reconstructing the history of a personal library or a history of reading based on only traces of reading, often not even in the original copies of the books themselves. If this can be considered a form of library history, this article has suggested an answer to the question how an interdisciplinary combination of library history and genetic criticism can proceed in the case of Joyce and Beckett. Both Joyce and Beckett wrote complex books that contain numerous allusions. Digital library history can enable readers both to discover such allusions and to further explore the intertextual complexity of these novels to the full.

28 According to the TEI guidelines, ‘The value unknown should be used in cases where the encoder does not wish to assert an opinion about the matter.’ (www.tei-c.org/Guidelines).