Text Encoding for Digital Scholarly Editions

Lecture

BSSDH | Digital Humanities in Latvia | 24 July 2019
Wout Dillen
University of Antwerp

@WoutDLN
The Digital Scholarly Edition
The Digital Scholarly Edition
Proust, Marcel: À la recherche du temps perdu: Du Côté de chez Swann

Paris: Editions de la Nouvelle Revue Française, 1923
Inscription: 'S. B. Beckett'
Inscription: 'P / Samuel Beckett // 254,041'
Inscription: 'For / Reading University Library / Samuel Beckett / april 1973'

 « Mais elle sera ravie de descendre ! Cela lui sera beaucoup plus de plaisir de causer avec vous que de s’ennuyer là-haut. » Hélas ! Swann en avait fait l’expérience, les bonnes intentions d’un tiers sont sans pouvoir sur une femme qui s’improvise de sentir pourvu qu’elle ait dans une fête par quelqu’un qu’elle aime pas. Souvent, l’ami redescend seul.

Ma mère ne vint pas, et sans ménagements pour mon amour-propre (engagé à ce que la fabule de la recherche dont elle était constamment en lisière de lui dire le résultat ne fût pas démentie) me fit dire par Françoise ces mots : « Il n’y a pas de réponse que depuis j’ai si souvent entendu des concierges de palais ou des valets de pied de tapis, rapporter à quelque pauvre fille qui s’étonne : « Comment, il n’a rien dit, mais c’est impossible ! Vous avez pourtant bien remis ma lettre. C’est bien, je vais attendre encore. » Et — de même qu’elle assure invariablement n’avoir pas besoin du bec supplémentaire que le concierge veuille allumer pour elle, et reste là, n’entendant plus que les rares propos sur le temps qu’il fait échanger entre le congéreur et la plaignante, elle dénonce les fées de la réalité de la réalité en apportant à la croyance la croyance du temps, et reste là, n’entendant plus que les rares propos sur le temps qu’il fait échanger entre le congéreur et la plaignante, elle dénonce les fées de la réalité de la réalité en apportant à la croyance la croyance du temps.
12 * Honour thy father and thy mother, that thy days may be long upon the land which the LORD thy God giueth thee.

13 * Thou shalt not kill.

14 Thou shalt not commit adultery.

15 Thou shalt not steal.

16 Thou shalt not bear false witness against thy neighbour.

17 * Thou shalt not covet thy neighbour’s man-servant, nor his maid-servant, nor his ox, nor his ass, nor anything that is thy neighbour’s.
‘A re-drawn and simplified version of the stemma published in Maas (1960, 7)’

Designed and published for the ‘Archetype’ definition in the ‘Parvum lexicon stemmatologicum’.
The Digital Scholarly Edition

Textual Criticism:

- What is the (closest) authentic text of the work?
- How was this text transmitted and transformed over time?
- How does this knowledge influence our interpretation of the work?
12 * Honour thy father and thy mother, that thy days may be long upon the land which the Lord thy God giueth thee.

13 * Thou shalt not kill.

14 Thou shalt not commit adultery.

15 Thou shalt not Steale.

16 Thou shalt not bear false wit against thy neighbour.

17 * Thou shalt not covet thy neighbour's house, thou shalt not covet thy neighbour's man-servant, nor his maid-servant, nor...
The Digital **Scholarly Edition**

Textual Criticism:

- What is the (closest) authentic text of the work?
- How was this text transmitted and transformed over time?
- How does this knowledge influence our interpretation of the work?

→ **Scholarly Edition:**

- Academic version of the work
- Contextualized with research of textual criticism
INTERESTING
A SANDWICH, WITH WORDS
IN CONGRESS, JULY 4, 1776.

The unanimous Declaration of the thirteen united States of America.

When in the Course of human Events, it becomes necessary for one People to dissolve the Political Bands which have connected them with another, and to avow their Independence, and to assume among the Family of Nations, the Station to which they are Called, it is the Duty of Nations to recognize and propagate the Truth. The history of the present King of Great Britain is a history of repeated injuries and usurpations, all having in direct Object the establishment of an absolute Tyranny over these States.

He has refused his Assent to Laws, the most wholesome and necessary for the public good.

He has dissolved Representative Houses repeatedly for opposing with manly Firmness their Invasion by着他.

He has refused to pass other Laws for the accommodation of large districts of People, until such districts were reduced under direct control. He has refused for a long time after such Laws were passed, to cause such Laws to be erected, under the Legislative Power, incapable of Combination; and has, by various indirect Methods, weakened or confounded the Power of the People.

He has plundered our seas, ravaged our Coasts, burnt our towns, and destroyed the lives of our people.

He has trivialized his Sacraments, parodying our public proceedings, and by various means, protecting his own rights and avert the force of our laws and the Punishment for their infraction.

He has excited domestic insurrections among us, and has endeavored to bring on the inhabitants of our frontiers the merciless Indian Savages, whose rage for revenge cannot be quenched, but only excited and愈次.

He has plundered our seas, ravaged our Coasts, burnt our towns, and destroyed the lives of our people.

He has trivialized his Sacraments, parodying our public proceedings, and by various means, protecting his own rights and avert the force of our laws and the Punishment for their infraction.

He has excited domestic insurrections among us, and has endeavored to bring on the inhabitants of our frontiers the merciless Indian Savages, whose rage for revenge cannot be quenched, but only excited and愈次.
A Declaration by the Representatives of the United States of America, in Congress assembled.

When in the course of human events it becomes necessary for one people to dissolve the political bands which have connected them with another, and to assume among the powers of the earth the separate and equal station to which the laws of nature and of nature’s God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to this separation.

We hold these truths to be self-evident, that all men are created equal, independent, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the pursuit of Happiness. That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed.
As the sentiments of men are known not only by what they receive, but [by] what they reject also, I will state the form of the Declaration as originally reported. The parts struck out by Congress shall be distinguished by a black line drawn under them, and those inserted by them shall be placed in the margin, or in a concurrent column.
Autobiography by Thomas Jefferson

The pusillanimous idea that we had friends in England worth keeping terms with, still haunted the minds of many.
The history of the present king of Great Britain is a history of unremitting injuries & usurpations, among which appears no solitary fact to contradict the uniform tenor of the rest but all have in direct object the establishment of an absolute tyranny over these states.
Autobiography by Thomas Jefferson

This piratical warfare, the opprobrium of INFIDEL powers, is the warfare of the CHRISTIAN king of Great Britain. Determined to keep open a market where MEN should be bought & sold, he has prostituted his negative for suppressing every legislative attempt to prohibit or to restrain this execrable commerce.
he has endeavored to bring on the inhabitants of our frontiers the merciless Indian savages, whose known rule of warfare is an undistinguished destruction of all ages, sexes, and conditions of existence.

He has incited treacherous insurrections of our fellow-citizens, with the allurements of forswearing and confiscation of our property.

he has constrained our fellow citizens for the support of his unrighteous war, to become the executioners of their friends and Brethren, he has waged cruel war against human nature itself, violating it’s most sacred rights of life and liberty in the persons of a distant people who never offend him, captivating & carrying them into slavery in another hemisphere, or to incur miserable death in their transportation thither. the
How the hidden word ‘subjects’ was discovered

1. Experts at the Library of Congress used a technique called hyperspectral imaging. The light reflected on the document is captured by the spectrometer optics.

2. Inside the device, light is redirected toward an optical dispersing element such as a prism that splits the light into narrow wavelength bands.

3. Hundreds of detectors process each of the separated wavelengths, transforming them into electronic signals that are interpreted by imaging equipment. Researchers also utilized other analysis software to enhance the underlying text.

SOURCE: Library of Congress | Alberto Cuadra and Karen Yourish
The Washington Post - July 3, 2010

http://www.washingtonpost.com/wp-dyn/content/graphic/2010/07/02/GR2010070205779.html
our fellow citizens,
The Digital Scholarly Edition

Textual Criticism:

- What is the (closest) authentic text of the work?
- How was this text transmitted and transformed over time?
- How does this knowledge influence our interpretation of the work?

→ Scholarly Edition:

- Academic version of the work
- Contextualized with research of textual criticism
The Digital Scholarly Edition

Textual Criticism:

- What is the (closest) authentic text of the work?
- How was this text transmitted and transformed over time?
- How does this knowledge influence our interpretation of the work?

→ Scholarly Edition:

- Academic version of the work
- Contextualized with research of textual criticism
The Digital **Scholarly Edition**

What is **authentic**? What do we use as the **base text**?

- The first edition?
- The last approved edition?
- Are drafts and proofs more or less important than published versions?
- What to do with copy-editing?
- What to do with the bibliographical aspects of the book?

→ Different ‘schools’ of textual criticism
The Digital Scholarly Edition

German Editionswissenschaft

- Document (Historical Critical Edition)

Anglo-American Scholarly Editing

- Text (Copy-Text Edition)

French Critique Génétique

- Writing Process (Genetic Edition)
Textual Criticism and (Digital) Scholarly Editing

Last decades: rappochement

- Three monolingual discussions become one multilingual one
- There is more to Textual Criticism than just these three traditions!
The Digital Scholarly Edition

Lexicon of Scholarly Editing
A Multilingual Lexicon for a Multilingual Discipline - #LexiconSE (beta)

http://uahost.uantwerpen.be/lse

#LexiconSE
Lexicon of Scholarly Editing
A Multilingual Lexicon for a Multilingual Discipline - #LexiconSE (beta)

Home  Lexicon  Bibliography  Zotero  News  Credits  Subscribe  Contribute  Contact  Affiliations & Links

Home

A Lexicon of Scholarly Editing

The Babel of Scholarly Editing

In November 2009, the topic of the annual conference of the European Society for Textual Scholarship was *Texts beyond Borders: Multilingualism and Textual Scholarship* (19-21 November 2009). The logo of the conference was Peter Brueghel the Elder's image of the Tower of Babel, the so-called little version (kept at the Museum Boijmans Van Beuningen in Rotterdam).

At this and following conferences, the need was expressed by several members of ESTS to create a lexicon of scholarly editing similar to undertakings in different disciplines and editorial traditions, and in different linguistic areas, a nice example being the French *Dictionnaire de critique génétique* (edited by Daniel Ferrer, Lydie Rauzier and Aurélie Crasson). Given the divergence of traditions, languages and contexts, such an undertaking is almost ‘doomed to fail’ from the start – to quote Samuel Beckett.

Towards a Lexicon of Scholarly Editing

But Beckett is also the author of that other quote: ‘Fall better.’ Under this motto, this lexicon was initiated in November 2012 by the European Society for Textual Scholarship (ESTS) and the Centre for Manuscript Genetics (part of the research group Literature and Modernity, University of Antwerp), as part of ESTS’s aim to provide an international and interdisciplinary forum for the theory and practice of textual scholarship in Europe.

Its aim is not to create new definitions, but to gather existing definitions for every entry in the lexicon. Many of these concepts have given rise to lively debates in the past and several eminent scholars have made courageous attempts to define them in monographs or scholarly journals. Every entry will be defined by means of one or more quotations from giants, offering their shoulders to future generations to stand on.

All quotations are referenced and the text from which they derive can be found in the bibliography. Although the lingua franca of the ESTS and of this English is English, the definitions do not need to be

PROGRESS

Entries: 420
Definitions: 888
Sources: 125
Languages: 7 (EN, DE, FR, IT, LA, NL, SP)
(last update: 13 June 2016)

CATEGORIES

❯ Author
❯ Basic Terminology
❯ Document
❯ Editions
❯ Editor
❯ Genetic Criticism
❯ Markup
❯ Palaeography
❯ Revision
❯ Scholarly Digital Editing
❯ Text

RECENT NEWS

❯ The Lexicon at Palermo
❯ The Lexicon is Going to Leicestersh - CIP ESTS 2015
❯ CIP: DI/TrBenelux 2015
❯ CIP: ‘Blue Skies Above, Solid Ground Below’
❯ Dutch Definitions in the Lexicon

USER PORTAL

Log in
Lexicon of Scholarly Editing
A Multilingual Lexicon for a Multilingual Discipline - #LexiconSE (beta)

Showing results for: document

ON THE MATERIALITY OF TEXT

...the scenes of The Book she is currently writing. As her post shows, she is working on concepts like text, work, document, and on the emphasis different types of scholars put on materiality and physical documents in their research. These topics are close to my own research, as I’m finishing up the introduction to my thesis, and know that I’ll have to be very clear in indicating which definitions of these highly contested concepts I’ll use.

DOCUMENT

document—the physical vessel (such as a book, manuscript, phonograph record, computer tape) that contains the text (Shillingsburg 1986, 170). document. The physical vessel (such as a book, manuscript, phonograph record, computer tape) that contains (or incarnates) the text (Shillingsburg 1996, 174). Een document kan dan gedefinieerd worden als het fysieke medium...

DOCUMENT (MULTI-VERSION)
The Multi-Version Document or MVD mode represents all the versions of a work, whether they arise from corrections to a text or from the copying of one original text into several variant versions, or some combination of the two, as four atomic operations: insertion, deletion, substitution, and transposition (Schmidt and ...)

DOCUMENT TYPE

The grammar defining how elements may legally be combined in a particular class of texts is known as a document type (Burnard 1991, section 2).

TEXT (DOCUMENT)

Der Text ist hier nicht nur Zeichenkette, sondern ein mehrdimensionaler Schriftraum mit vielfältigen Ausdrucksrepertoire. Neben dem (linearen) "linguistischen text" gibt es einen (nicht-linearen)"...
**Document**

By Elias Nury | 13 December 2014 | Comments Off on document |

A document is the physical vessel (such as a book, manuscript, phonograph record, computer tape) that contains the text (Shillingsburg 1999, 170).

A document. The physical vessel (such as a book, manuscript, phonograph record, computer tape) that contains (or incarnates) the text (Shillingsburg 1996, 174).

Een document kan dan gedefinieerd worden als het *fysieke materiaal* (meestal *papier* en *inkt*) waarop de configuratie van tekens is aangebracht die samen de *tekst* vormen (Van Hulle 1998, 53).


**RELATED ENTRIES**

- record
- text (document)
- text carrier
- manuscript
- witness
- transcription
- editing (documentary)
- orientation (documentary)
The Digital Scholarly Edition

Textual Criticism:
a research field concerned with the way text is transmitted over time, and the way in which this transmission may inform our reading of the text.

Scholarly Edition:
The field’s output format, aiming to represent an authoritative, edited text that takes that process of textual transmission into account in some way or other.
The Digital Scholarly Edition
The **Digital** Scholarly Edition
En toen gebeurde dergens, wat kon hem een moment van glijdend, bovenheid of more brakke. Quintus in september in twijfel was aan zijn, firmelik door de brievenmando en de bloemen en gegoede, che rubenendens terug te halen naar het kasteel. "Quintus zal altijd voetgangers naar het pijnlijke, brekende een benige scherpevlekk, een 'olifant.'"
Hypertext is the combination of natural-language with the computer’s capacities for interactive, branching or dynamic display, *when explicitly used as a medium*. Or, to define it more broadly, ‘hypertext’ is the generic term for any text which cannot be printed (or printed conveniently) on a conventional page, or used conveniently when bound between conventional covers. ‘Non-linear text’ might be a fair approximation. (Nelson 1967, 195).
2. quo] a quo Wes. fort. recte.
satis scite] HDF; satis seis e M.
mane] M; manere HF.

Ep. 654 (Att. xiii. 47a).

1. malui] M; malim alii.
moleste færem] moleste ferre M; et moleste ferre Wes.

Ep. 655 (Fam. xvi. 19).

suo] om. D et Index MH.
potest] M; potes H (sed una littera erasa); potes DF.
nihil] om. HF.

Ep. 656 (Att. xiii. 48).

scripsisse] Lamb.; scripsisti M.
2. iuva] via M\textsuperscript{1} idemque infra.
putato] C; puto M.
Asturam] Wes.; ad sturae M.
si] Zb; om. M.

Ep. 659 (Att. xiii. 39).

sibi] v. c. Vict.; tibi M.
\textgreek{theo\tilde{\nu}}] Vict.; \textgreek{O\mathrm{\on}} M.
\textgreek{P\lambda\lambda\dot{a}o\v{s}}] \textgreek{p\epsilon\textgreek{r}i P\lambda\lambda\dot{a}dos} Orelli coll.
Nat. Deor. i. 41; \textgreek{'Apollod\omega{r}ou} Hirzel;
pant\delta\sigma\varsigma Gurlitt; vide Comm.

Ep. 660 (Att. xiii. 40).

1. autem\textsuperscript{2} Tu, \textit{futilem est}, nos
Hypertext is the combination of natural-language with the computer’s capacities for interactive, branching or dynamic display, *when explicitly used as a medium*. Or, to define it more broadly, ‘hypertext’ is the generic term for any text which cannot be printed (or printed *conveniently*) on a *conventional* page, or used *conveniently* when bound between *conventional* covers. ‘Non-linear text’ might be a fair approximation.

Nelson 1967, 195
The **Digital** Scholarly Edition

Digital scholarly editions are not just scholarly editions in digital media. I distinguish between *digital* and *digitized*. A digitized print edition is not a ‘digital edition’ in the strict sense used here. A digital edition can not be printed without a loss of information and/or functionality. **The digital edition is guided by a different paradigm.** If the paradigm of an edition is limited to the two-dimensional space of the ‘page’ and to typographic means of information representation, than it’s not a digital edition.

Sahle 2008 (online)
The **Digital** Scholarly Edition

In other words: a Digital Scholarly Edition tries to take full advantage of the digital medium, e.g. by focussing on:

- Providing an intuitive reading experience for comparing different versions
- Satisfying the needs of different ‘schools’ of textual scholarship simultaneously (different visualizations on the basis of the same data)
- Allowing the user to query the data in new user-friendly, flexible and unforeseen ways (GUI / API)
- Engaging users to contribute to the edition’s content in new ways
- Achieving new levels of accountability by encoding interpretative and editorial decisions on the micro level
- ...

WELCOME

The Beckett Digital Manuscript Project is a collaboration between the Centre for Manuscript Genetics (University of Antwerp), the Beckett International Foundation (University of Reading) and the Harry Ransom Humanities Research Center (University of Texas at Austin), with the kind permission of the Estate of Samuel Beckett.

The purpose of the Beckett Digital Manuscript Project is to reunite the manuscripts of Samuel Beckett's works in a digital way; and to facilitate genetic research: the project brings together digital facsimiles of documents that are now preserved in different holding libraries, and adds transcriptions of Beckett's manuscripts: tools for bilingual and genetic version comparison, a search engine, and an analysis of the textual genesis of his works.

BECKETT DIGITAL LIBRARY ADDED

We are delighted to announce the publication of the Beckett Digital Library. The accompanying monograph Samuel Beckett's Library, written by Dirk Van Hulle and Mark Nixon, is available through Cambridge University Press.

KRAPPS LAST TAPE / LA DEMANDE BANDE

We are delighted to announce the publication of our third genetic edition: Krappp's Last Tape / La Demande Bande. The accompanying monograph on its genesis is available in Belgium through University Press Antwerp, and internationally through Bloomsbury Academic.

© 2016 Samuel Beckett Digital Manuscript Project

Directors: Dirk Van Hulle and Mark Nixon | Technical Realisation: Vincent Neyt

Under the auspices of the Centre for Manuscript Genetics (University of Antwerp), the Beckett International Foundation (University of Reading), the Harry Ransom Humanities Research Center (Austin, Texas) and the Estate of Samuel Beckett.

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 313609.

Stirrings Still / Souffleuses et Comment dire / what is the word © Samuel Beckett 1988, 1989 and the Estate of Samuel Beckett. The right of Samuel Beckett to be identified as the author of this work has been asserted in accordance with Section 77 of the Copyright, Designs and Patents Act 1988.

L'Innumérable / The Unnumerable © Samuel Beckett 1953 and the Estates of Samuel Beckett.

www.beckettarchive.org
Welcome

The Beckett Digital Manuscript Project is a collaboration between the Centre for Manuscript Genetics (University of Antwerp), the Beckett International Foundation (University of Reading) and the Harry Ransom Humanities Research Center (University of Texas at Austin), with the kind permission of the Estate of Samuel Beckett.

The purpose of the Beckett Digital Manuscript Project is to reunite the manuscripts of Samuel Beckett’s works in a digital way, and to facilitate genetic research: the project brings together digital facsimiles of documents that are now preserved in different holding libraries, and adds transcriptions of Beckett’s manuscripts, tools for bilingual and genetic version comparison, a search engine, and an analysis of the textual genesis of his works.

MLA Prize for a Bibliography, Archive, or Digital Project

On 6 December 2018, the Modern Language Association of America announced the BDMP as the winner of the eleventh Modern Language Association Prize for a Bibliography, Archive, or Digital Project.

Read more on our News & Updates page.

Samuel Beckett: A Bibliography

We are honoured to announce the publication of the first part of Bron Mitchell’s Samuel Beckett: A Bibliography: Part I: The Early Years, 1906–1953.

Beckett Digital Library (BDL)

We are delighted to announce the publication of the Beckett Digital Library. The accompanying monograph, Samuel Beckett’s Library, written by Dirk Van Hulle and Mark Nixon is available through Cambridge University Press.

© 2019 Samuel Beckett Digital Manuscript Project

Directors: Dirk Van Hulle and Mark Nixon | Technical realisation: Vincent Nayt

Under the auspices of the Centre for Manuscript Genetics (University of Antwerp), the Beckett International Foundation (University of Reading), the Harry Ransom Humanities Research Center (Austin, Texas) and the Estate of Samuel Beckett.

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n°313099.

Strings Still / Sous la plume et Comment dire / what is the word? © Samuel Beckett 1988, 1989 and the Estate of Samuel Beckett. The right of Samuel Beckett to be identified as the author of this work has been asserted in accordance with Section 77 of the Copyright, Designs and Patents Act 1988.

L’Innommable / The Unnamable © Samuel Beckett 1953 and the Estate of Samuel Beckett.

Krap’s Last Tape / La Dernière Bande © Samuel Beckett 1958 and the Estate of Samuel Beckett.

Molloy © Samuel Beckett 1951 and the Estate of Samuel Beckett.


Fin de partie / Endgame © Samuel Beckett 1957 and the Estate of Samuel Beckett.

No part of this publication may be reproduced in any form or by any electronic or mechanical means without permission in writing from the publisher and from the Estate of Samuel Beckett.
Digitization
Digitization

Putting together a ‘Genetic Dossier’.

Loose notes
Digitization

Putting together a ‘Genetic Dossier’.
Digitization

Putting together a ‘Genetic Dossier’.

Early Drafts
Digitization

Putting together a ‘Genetic Dossier’.

Typecripts

A late evening in the nineteen eighties.

Kropp’s dam.

Front centre, a small table, the two drawers of which open towards audience.

Sitting at the table facing front, i.e. across from the drawers, a weary old man: Kropp.

Rusty black narrow trousers too small for him. Rusty black sleeveless waistcoat, four capacious pockets. Heavy silver watch and chain. Gray white shirt open at neck, no collar.

Surprising pair of dirty white boots, slim too, very narrow and pointed.


Very near-sighted (but unaccommodated) hand of hearing.

Cracked voice. Sibilant accent on consonants.

Walking on air.

On the table a tape-recorder with microphone and a number of cardboard boxes containing reels of recorded tapes.

Table and immediately adjacent area in strong white light.

Seat of stage in darkness.

Kropp remains a moment motionless, looks at his watch. Fumbles in his waistcoat pockets, takes out an envelope, puts it back. Fumbles, takes out a small bunch of keys, raises it to his eyes, chooses a key, gets up and moves to front of table. He stoops, unlocks, looks in, peers into it, feels about inside it, takes out a rectangle, it back, slopes and locks drawer, unlocks and opens drawer, peers into it, feels about inside it, takes out a long banana, chooses and locks drawer, puts key back in his pocket. He turns, advances to edge of stage, halts, and banana, peels it, drops skin at his feet, puts end of banana in his mouth and remains motionless, staring vacuously before him. Finally he bites off the end, turns aside and begins pacing to and fro at edge of stage, in the light, i.e. not more than four or five paces either way, meditatively eating banana. He eats on skin, slides, nearly falls, recovers himself, stoops and peels at skin and finally pushes it, still eating, with his foot over the edge of stage into pit. He resumes his pacing, finishes banana, returns to table, sits down, remains a moment motionless, takes keys from his pocket, raises them to his face, chooses a key, gets up and moves to front of table, opens drawer, takes out a second banana, chooses and locks drawer, puts key back in his pocket, turns and advances to edge of stage, halts, banana, peels it, drops skin at his feet, puts end of banana in his mouth and remains motionless, staring vacuously before him.
Digitisation

Scanning the materials: some quality measures:

- Minimum 300 dpi
- 24 bit
- lossless format (TIFF)
- Color and size references
- Scanning the document, not just the text (plenty of margins)
Transcription (XML)
Markup

The term markup is derived from the traditional publishing practice of **marking up** a manuscript, which involves adding handwritten **annotations** in the margins and text of a document.
Markup

For a computer to be able to distinguish between our annotations and the actual text, it is important that we again use a series of conventions regarding the way we format our annotations. That we have rules to keep them “syntactically” distinguishable from one another.

For this we use Markup Languages, and specifically in this context XML (eXtensible Markup Language).
The **Digital** Scholarly Edition

Example: Hello **World**!

Text

Hello World!
The **Digital** Scholarly Edition

Example: Hello *World!*

Tags: mark start and end

```
<title>Hello World!</title>
```
The Digital Scholarly Edition

Example: Hello World!

Attribute: specify further with value

<title type="blog">Hello World!</title>
Example: Hello World!

Nesting

<title>Hello World!</title>
The Digital Scholarly Edition

Types of markup:

- punctuational markup
- presentational markup
- procedural markup
- descriptive markup
- referential markup
- metamarkup

(Coombs, Rengaer, DeRose 1987, 935-937).
What is Text, Really?

Steven J. DeRose
praxis
David G. Durand
Brandeis University
Ellis Mylonas
Harvard University
Allen H. Renear
Brown University

ABSTRACT

The way in which text is represented on a computer affects the kinds of uses to which it can be put by its creator and by subsequent users. The electronic document model currently in use is impoverished and restrictive. The authors argue that text is best represented as an ordered hierarchy of content objects (OHCO), because that is what text really is. This model conforms with emerging standards such as SGML and contains within it advantages for the writer, publisher, and researcher. The authors then describe how the hierarchical model can allow future use and reuse of the document as a database, hypertext, or network.

Although computing visionaries have spoken of an almost magical future world of accessible information and communication, years of progress have still failed to realize Alan Kay's image of "piloting a jet plane through information space." The authors emphatically believe that the computer has immense potential to assist people in general, and scholars in particular, in their intellectual work. This assistance will extend across such activities as researching, thinking, and writing, as well as recording and organizing informa-
Text as an Ordered Hierarchy of Content Objects (OHCO)

<book>

  <chapter>CHAPTER 1</chapter>

  <paragraph>

    <sentence>A sentence.</sentence>

    <sentence>Another sentence</sentence>

  </paragraph>

</book>
Overlapping hierarchies:

One of Shakespeare’s most famous quotes is probably “To be or not to be, that is the question’.

One of Shakespeare’s most famous <u>quotes</u> is probably <quote>“To be or not to be, </quote>that is the question’.</quote>
The Digital Scholarly Edition

Overlapping hierarchies:

One of Shakespeare’s most famous quotes is probably “To be or not to be, that is the question’.

One of Shakespeare’s most famous <u>quotes is probably <quote>“To be or not to be,</quote> that is the question’.</quote>
The Digital Scholarly Edition

Overlapping hierarchies:

One of Shakespeare’s most famous quotes is probably “To be or not to be, that is the question’.

One of Shakespeare’s most famous <u>quotes is probably</u> <quote><u>“To be or not to be, </u>that is the question’.</quote>
Samuel Beckett’s *Malone Dies* is a great novel.
Samuel Beckett’s *Malone Dies* is a *great* novel.
The Digital Scholarly Edition

Example: Hello World!

Underline

<title>Hello <u>World</u>!</title>
Example: Hello World!

Underline

<title>Hello <underlined>World</underlined>!</title>
The **Digital Scholarly Edition**

It is of the utmost importance that electronic scholarly editions adhere to **international standards**. Projects that are idiosyncratic are almost certain to remain stand-alone efforts: they have effectively abandoned the possibility of interoperability.

(Price 2007, 442).
The Digital Scholarly Edition

TEI

TEXT ENCODING INITIATIVE
Welcome to the BDMP Encoding Manual!

This website contains the official and most up-to-date version of the Encoding Manual used by the Beckett Digital Manuscript Project (BDMP) for encoding Samuel Beckett's manuscripts in TEI-compliant XML. The BDMP offers a genetic edition of Samuel Beckett's works, and falls under the auspices of the Centre for Manuscript Genetics at the University of Antwerp, the Beckett International Foundation at the University of Reading, the Harry Ransom Research Center at the University of Texas at Austin, and the Estate of Samuel Beckett.

The primary purpose of this website is to serve as an encoding resource for our collaborators. In our General Remarks section, they can refresh their memories concerning the basics of TEI-XML encoding in general, and of our project-specific encoding guidelines in particular. And in our section on The Tags, they will find all the information they need on the tags we use to encode both The Header and The Text. But because we think this information may also be useful for other researchers or general public alike, we have added an About us section and a Contact section to take the BDMP and/or genetic criticism to take a look under the hood of the BDMP, and see how we encode our transcriptions. For more information, please check out our about page. For the most recent update reports, please consult our News and Updates blog. And for questions, comment, bug reports, etc., please refer to our Contact page.
Après le premier acte, on se résout à continuer l'action sans la couper.
arrêter. Afin de parler. On se met à parler en le voulant comme si l’on pouvait s’arrêter à volonté.

<del>arrêter.</del> Afin de parler. On se met à parler <lb/> comme si l’on pouvait s’arrêter <del>à volenté</del><add>en le voulant</add><lb/>
arrêter. Afin de parler. On se met à parler en le voulant comme si l’on pouvait s’arrêter à volonté.

<del>arrêter.</del> Afin de parler. On se met à parler <lb/> comme si l’on pouvait s’arrêter <del>à volonté</del><add>en le voulant</add><lb/>

(Beckett 2013, MS-HRC-SB-3-10 p.08v)
arrêté. Afin de parler. On se met à parler en le voulant comme si l’on pouvait s’arrêter à volonté.

<del>arrêter.</del> Afin de parler. On se met à parler <lb/> comme si l’on pouvait s’arrêter <del>à volonté</del><add>en le voulant</add><lb/>

(Beckett 2013, MS-HRC-SB-3-10 p.08v)
arrêter. Afin de parler. On se met à parler en le voulant comme si l’on pouvait s’arrêter à volonté.

<del>arrêter.</del> Afin de parler. On se met à parler <lb/> comme si l’on pouvait s’arrêter <del>à volonté</del><add>en le voulant</add><lb/>

(Beckett 2013, MS-HRC-SB-3-10 p.08v)
arrêter. Afin de parler. On se met à parler en le voulant comme si l'on pouvait s’arrêter à volonté.

<del>arrêter.</del> Afin de parler. On se met à parler <lb/> comme si l'on pouvait s’arrêter <del>à volonté</del><add>en le voulant</add><lb/>

(Beckett 2013, MS-HRC-SB-3-10 p.08v)
Which editorial practice will serve my text best?

➤ How can I best exploit the full potential of my text?
License

This work is licensed under a Creative Commons Attribution 4.0 International Public License.

All works of other authors cited, linked, and referred to here are their intellectual property and are used for academic purposes only.