

# Encoding, Visualizing, and Generating Variation in Fernando Pessoa's *Livro do Desassossego*

Manuel Portela and António Rito Silva

*Abstract:* One of the goals of the *LdoD Digital Archive* is to show Pessoa's *Book of Disquiet* as a network of potential authorial intentions and a conjectural construction of its successive editors. Our digital representation of the dynamics of textual and bibliographical variation depends on both XML encoding of variation sites (deletions, additions, substitutions, etc.) and metatextual information concerning authorial and editorial witnesses (date, order, heteronym, etc.). While TEI-XML markup may be considered as a particular kind of critical apparatus on its own, it is through visualization tools and graphical interface that users will be able to engage critically with the dynamics of variation in authorial and editorial witnesses. Besides representation of textual and bibliographical variation in the work's genetic and editorial history, the *LdoD Digital Archive* will enable readers to generate new virtual forms of the work, assuming the roles of editor and/or author. Our paper discusses the theoretical and technical aspects of the various strategies adopted for encoding, visualizing, and generating variation in the *LdoD Digital Archive*. These issues are discussed with the support of a prototype currently under development. *Keywords:* Fernando Pessoa, digital archive, social edition, TEI-encoding, visualization, variants, variation

FERNANDO PESSOA'S *Livro do Desassossego* (*Book of Disquiet*) is an unfinished book project.<sup>1</sup> Pessoa wrote more than five hundred texts meant for this work between 1913 and 1935, the year of his

---

<sup>1</sup> "Nenhum Problema Tem Solução: Um Arquivo Digital do *Livro do Desassossego*" [No Problem Has a Solution: A Digital Archive of the *Book of Disquiet*] is a research project of the Centre for Portuguese Literature at the University of Coimbra (2012–2015), funded by FCT (Foundation for Science and Technology). Principal investigator: Manuel Portela. Reference: PTDC/CLE-LLI/118713/2010. Co-funded by FEDER (European Regional Development Fund), through Axis 1 of the Operational Competitiveness Program (POFC) of the National Strategic Framework (QREN). COMPETE: FCOMP-01-0124-FEDER-019715.

death.<sup>2</sup> The first edition of this book was published only in 1982,<sup>3</sup> and another three major versions have been published since then (1990 [7<sup>th</sup> revised edition, 2013], 1998 [11<sup>th</sup> revised edition, 2014], 2010 [3<sup>rd</sup> revised edition, 2014]). As it exists today, the *Livro do Desassossego* may be characterized as (1) a set of autograph (manuscript and typescript) fragments, (2) mostly unpublished at the time of Pessoa's death, which have been (3) transcribed, selected, and organized into four different editions, implying (4) various critical and genetic interpretations of what constitutes this book. Editions show four major types of variation: variation in readings of particular passages, in selection of fragments, in their ordering, and also in heteronym attribution.<sup>4</sup>

Those editorial instantiations have given material expression to four different models of constructing the *Livro*. We could

---

<sup>2</sup> Currently, the Portuguese National Library has catalogued 722 sheets as belonging to the *Livro do Desassossego*, of which 374 are typescripts, while 348 are manuscripts. Some of them are written on recto and verso. Only 12 texts from the *Livro* were published by Pessoa. Texts explicitly assigned by Pessoa to *Livro do Desassossego* contain the annotation "L. do D." However, there are more than two hundred texts without the "L. do D." annotation that also belong (or have been ascribed by editors as belonging) to the *Livro*. The total set of fragments in each edition has varied either because new texts have been discovered in Pessoa's Archive, or because editors have decided to include or exclude particular texts. Another reason for variation originates in the fact that some documents have been interpreted as one single text or as more than one text. The number of fragments in the editions that we have encoded for the *LdoD Archive* is as follows: the edition by Jacinto do Prado Coelho (Pessoa 1982), 520 fragments; by Teresa Sobral Cunha (Pessoa 2008), 748 fragments; by Richard Zenith (Pessoa 2012), 514 fragments; and by Jerónimo Pizarro (Pessoa 2010), 586 fragments.

<sup>3</sup> This first edition, published in 1982, was edited by Jacinto do Prado Coelho and transcribed by Maria Aliete Galhoz and Teresa Sobral Cunha. Jorge de Sena had also started working on the preparation of an edition of the *Livro* in the 1960s. This work was never finished, although Sena wrote an introduction for the *Livro* in 1964. For a detailed editorial history of all the works by Fernando Pessoa, see Pizarro 2012, 29–90. For current views on *Livro do Desassossego*, see Pizarro 2013 and Zenith 2013.

<sup>4</sup> The first authorial persona for the *Livro* was Vicente Guedes, but the work was later reassigned by Pessoa to Bernardo Soares, a persona described by Pessoa as a "semi-heteronym". Although the authorial *personae* behind *Livro do Desassossego* tend not to be viewed as full heteronyms, heteronym attribution has been an important function in structuring the work. See also footnote 10 below.

summarize this history as follows: the first model orders fragments according to a combination of thematic and chronological proximity (Pessoa 1982; edited by Jacinto do Prado Coelho); the second model distinguishes between two periods of composition and their respective heteronyms (Vicente Guedes and Bernardo Soares), while strengthening the discursive unity of the fragments within each part, for example, by removing text numbering and by rearranging the internal structure of a certain number of more fragmentary texts (1990–2013; edited by Teresa Sobral Cunha); the third model considers the production of Bernardo Soares as the main axis of the work and anchors the remaining fragments so that Soares's voice becomes predominant, relegating the set of early large texts by Guedes to a final section (1998–2014; edited by Richard Zenith); finally, the fourth model produces a critical and genetic reconstruction based on the inferred chronology of the fragments, thus bringing the order of the *Livro* closer to its archival order (2010–2014; edited by Jerónimo Pizarro).<sup>5</sup>

One of the goals of the *Livro do Desassossego (LdoD) Digital Archive*, which is being developed at the Centre for Portuguese Literature (CLP) at the University of Coimbra, is to show Pessoa's book as a network of potential authorial intentions and a conjectural construction of its successive editors (Portela 2013). Our digital representation of the dynamics of textual and bibliographical variation depends on both XML encoding of variation sites (deletions, additions, substitutions, etc.) and metatextual information concerning authorial and editorial witnesses (date, order, heteronym attribution, etc.). While TEI-XML markup may be considered as a particular kind of critical apparatus on its own, it is through visualization tools and graphical interfaces that users will be able to engage critically with the dynamics of variation in authorial and editorial witnesses. Besides navigating through the archive's materials, users will be able to generate further textual and bibliographical variations by assuming the role of editors or even writers. This article discusses the theoretical and technical aspects of the various strategies adopted for encoding, visualizing and generating variation in the *LdoD Digital Archive*. Before presenting examples of

---

<sup>5</sup> For a recent analysis of the editions of *Livro do Desassossego*, see Silvestre 2014. Pessoa's editorial plans for the *Livro* have been closely examined by Sepúlveda 2013.

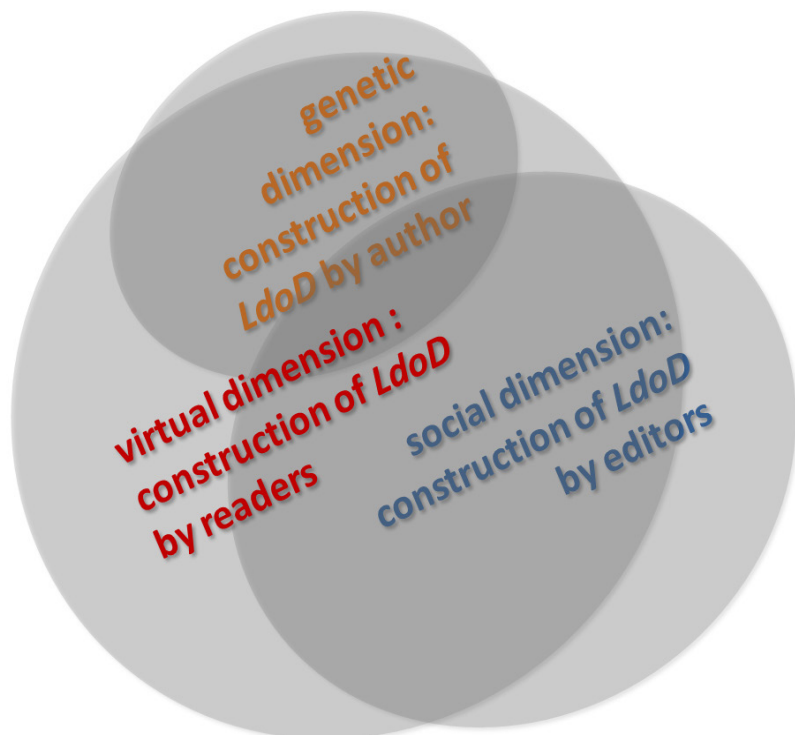


Figure 1: XML encoding and algorithmic processing of Fernando Pessoa's *Book of Disquiet* attempts to integrate a digital representation of autograph materials (genetic dimension) with a digital representation of editorial materials (social dimension), and both those levels with a digital simulation of writing and editing processes (virtual dimension).

those strategies, we will briefly explain the project rationale in its genetic, social and virtual dimensions.

### **From author's text to editor's text to reader's text**

One of the goals of the *LdoD* research project is to produce a comprehensive digital archive that integrates a genetic with a textual scholarship approach by treating both authorial witnesses and published editorial versions as elements in the work's textual and social history. Published versions are understood as variations, at both micro and macro-levels, on Pessoa's unfinished book project. A full integration of authorial and editorial witnesses is achieved

by treating all of them as actual versions of a number of possible versions of the *Book of Disquiet*. Thus authorial and editorial textual forms can be compared against each other at different scales — from the micro-level of word to the macro-level of book structure. This particular combination of transcription of authorial and editorial text enables readers and scholars not only to construct narratives of composition and narratives of editing and publication, but also to place each of those narratives in the context of the others. Our model for electronic encoding of textual fragments supports radial configurations that allow users to engage the work through three different facets: the genetic dimension; the social dimension; and the virtual dimension (Figure 1).

The virtual dimension enables readers to generate further variations by means of editing and writing acts. We propose to extend social editing theories to include the collaborative capabilities of the digital medium, suggesting that its participatory affordance can be used for generating new variations on the work's textual and bibliographical form (Silva and Portela 2013 and 2015). Thus "social editing" is conceived of in two different but related senses: one is derived from the theory of social editing, which describes the socialization of texts embodied in particular bibliographical codes in the historical archive; the other refers to social editing as a collaborative practice in web environments. This second sense may be understood as an extension of the first in digital media environments where literary processes can become a continuous process of production. The socialization of editing through Web 2.0 tools has specific implications, such as the possibility of continuous reediting — unlike the print medium, where the production of a new edition only happens when an earlier edition is no longer available or when there is enough market demand for several editions to compete amongst themselves.<sup>6</sup>

---

<sup>6</sup> Users of the archive will be able to register in the system and collaborate in the creation of virtual editions. This collaboration can take place informally but also in institutional contexts, particularly within schools and universities. We intend to set up several virtual communities of students and teachers who work with the *Book of Disquiet* at different locations, and support their use of the dynamic features of the Archive. Usability tests are being made to assist us in graphical interface design, and also for developing guidelines and tutorials that will explain the collaborative functions to future users. A second stage of project development (2016–2018) will be specifically concerned with

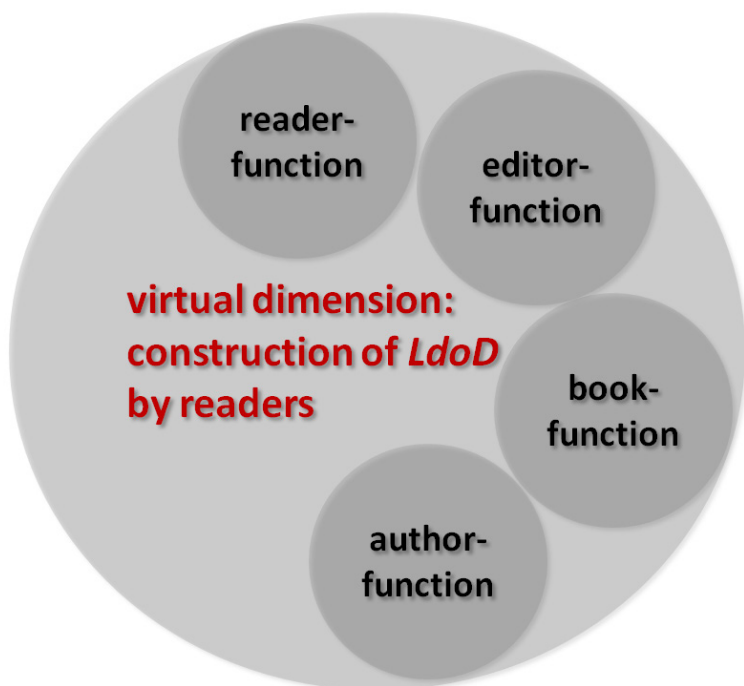


Figure 2: Readers can interact within the archive ecology according to four virtual roles

This programmatic virtualization of the *Book of Disquiet* is a performative intervention in the work's archive that takes place also at the level of editing and writing. Our model for electronic encoding of textual fragments is based on established principles in genetic criticism and textual scholarship informed by a desire to engage the fluidity and flexibility of our current media environment (Bryant 2002; Benel and Lejeune 2009; Fraistat and Jones 2009; Siemens et al. 2012). This model of the reconfigurative iterability of the work's authorial and editorial archive will enable users to virtualize Pessoa's book project according to four functions: reader-function, editor-function, book-function, and author-function (Figure 2; see also Portela and Silva 2014). The changing roles of users within the archive ecology enhance the dynamic character of our digital simulation of textuality as an ongoing material and interpretive

---

optimizing the collaborative affordances that we are programming.

process of textual transformation. By means of interaction between textual representation and textual transformation the *LdoD Archive* becomes an experimental engine for the simulation of literary processes.

### **The textual dynamics of the archive: Document, text, book, work**

In a recent article on theory of digital editions, Peter Robinson shows how printed editions have traditionally focused on the relationship of text to work, while digital editions have been more focused on the relationship of text to document. He suggests that “a scholarly edition must, so far as it can, illuminate both aspects of the text, both text-as-work and text-as-document” (Robinson 2013, 123). The *LdoD Archive* embodies a similar understanding of the nature of textual semiosis as process involving self and object in a continuous and co-dependent process of meaning production through acts of reading. Editing Pessoa’s centrifugal and reticular body of unpublished work is an especially acute experience of the productive function of reading in activating the force fields that allow you to move back and forth from document to text to book to work.

By placing digital facsimiles in the context of topographic transcriptions, the *LdoD Archive* enables users to experiment with the transit from document to text and from text to document. Situating both facsimile and topographic transcription in the context of the experts’ editions, the *LdoD Archive* shows four possible transitions from text to book and from book to text. To the extent that each text of each edition is contextualizable in an archive of authorial and editorial witnesses, it is the very process of construction of text from document and book from text that the genetic and social dimensions of the *LdoD Archive* place in evidence. The construction of the book — as the product either of a self-editing authorial act, or a series of posthumous editorial acts — becomes an instantiation of the conceptual and material process of identity and difference that enables text and book to emerge from a series of inscriptional marks and from the acts of reading and interpreting those marks.

The electronic edition of modern manuscripts that are unfinished book projects can be conceived according to different principles. The principle used by the editors of the *LdoD Archive* is based



on the unit “trecho” (passage; or, literally, stretch, i.e. a continuous piece of completed writing) or fragment (a continuous piece of writing in progress), understood as a certain textual extension with thematic or material evidence of textual unity, which can be further marked (or not) by graphical markers: for example, a larger space or a larger number of blank lines between two handwritten textual sequences suggests an interruption; or by genetic and editorial events, for example, a piece that has been typed or published in a magazine.<sup>7</sup> The units of composition included in the *LdoD Archive* are based on either the set of four critical editions, or the ensemble of digital facsimiles of authorial witnesses that correspond to that set.<sup>8</sup> On the other hand, it is possible to think of the act of writing as a speech act, that is, as a certain temporal unit of writing that does not always coincide with the documentary material unit (recto and verso of a loose leaf, or a set of contiguous pages in a notebook, for example) or with units of written discourse, such as the paragraph, or other units of bibliographical structure, such as the chapter. Pessoa’s writing practice for the *Livro* seems to emphasize this act of scripting as a unit of composition of the work itself.<sup>9</sup>

The existence of very short fragments, almost aphoristic in scale, along with fragments of varying length (one paragraph to one page to a few pages), written at very different moments in time, suggests precisely this noncoincidence between the temporal unit of

---

<sup>7</sup> The *LdoD Archive* uses the notion “fragment” to refer to all texts from the *Livro*. Although many texts can be considered fragmentary because they may have not been finished or revised, there are also many texts that are finished and revised pieces (“trechos”). Pessoa seems to have conceived the book as a particular arrangement of “trechos”. Regardless of their stage of completion and revision, all textual pieces may be considered textual fragments of the projected book.

<sup>8</sup> Although only Pizarro’s 2010 edition contains a detailed and extensive critical apparatus, we consider that Coelho’s, Sobral Cunha’s and Zenith’s minimal critical apparatus also qualifies them as critical editions (Pessoa 1982, 2012 and 2013a). The texts selected for transcription in the *LdoD Archive* are the total sum of all texts of those four scholarly editions, including those that have only been transcribed in appendices. There will be also a special section in the *LdoD Archive* with Pessoa’s notes and plans for the *Book of Disquiet*.

<sup>9</sup> This notion can be related to Peter Shillingsburg’s “script act” (2006), although what we want to emphasize here is the fact that most textual units in the *Livro* seem to coincide with one temporal unit of writing, i.e. one sustained period of continuous writing.



writing and the cumulative and retrospective process of accretion and rewriting that produces semantic coherence and syntactic cohesion. Each moment of writing (which is also a moment of self-consciousness of writing) originates a new self-contained thematic and stylistic unit. These units proliferate as fragments of a book in progress but they resist the material order of the book. Edward Vanhoutte (2006) characterizes the modern manuscript as a complex network of those temporal units of writing. Marta L. Werner (2011), in her turn, describes the modern manuscript as a record of the dynamics of text in the process of creating itself. She places it in a liminal space of private inscription which becomes physically reflected in its undisciplined textual condition.

By using the notion of “trecho” [brief textual section] to refer to the units of composition of *LdoD*, Fernando Pessoa is aware of this dimension of aggregation and sequencing of small textual units as one of the compositional principles of *LdoD*. The revision process that he imagines toward the bibliographical horizon seems to imply the simultaneous production of psychological coherence and stylistic consistency:

L. do D. (nota)

A organização do livro deve basear-se numa escolha, rígida quanto possível, dos trechos variadamente existentes, adaptando, porém, os mais antigos, que falhem á psychologia de Bernardo Soares, tal como agora surge, a essa vera psychologia. Aparte isso, ha que fazer uma revisão geral do próprio estylo, sem que ele perca, na expressão íntima, o devaneio e o desconnexo lógico que o caracterizam. Ha que estudar o caso de que se devem inserir trechos grandes, classificaveis sob titulos grandiosos, como a Marcha Funebre do Rei Luiz Segundo da Baviera, ou a Symphonia de uma Noite Inquieta. Ha a hypothese de deixar como está o trecho da Marcha Funebre, e ha a hypothese de a transferir para outro livro, em que ficassem os Grandes Trechos juntos. (Pessoa 1982, 8)

[L. do D. (Note)]

The organization of the book should be based on a choice, as rigid as possible, from the existing varying texts [“trechos”], adapting, however, the older ones, which may fail the psychology of Bernardo Soares, as it now appears, this true psychology. Apart from this, a general revision of his own style, without letting it lose, in

its intimate expression, the reverie and disjointed logic that define him. We must study the case whether to include the large texts, classifiable under grandiose titles such as *The Funeral March for King Ludwig the Second of Bavaria*, or *Symphony of a Restless Night*. There is the possibility of leaving the *Funeral March* as it stands, and there is the possibility of transferring it to another book, in which the *Large Texts* would stay together.]

Fernando Pessoa is thinking about submitting the fragments of the *Livro* to the conceptual and material coherence of the book form. He is recognizing both the disjointed dreamlike introspective style of Bernardo Soares, and also textual affinities among the large texts.<sup>10</sup> Bibliographical coherence seems to depend simultaneously on two separate logics: an external logic of organization that sequences and articulates its elements according to the syntactic structure and the horizon of codex totalization, which creates unity through its discrete and finite character; and an internal logic of organization that selects and associates fragments because of semantic and stylistic affinities, producing bibliographical unity through the cumulative effect of discursive coherence between associated brief and lengthy pieces. The difficulty in matching the material and discursive space of writing to the material and conceptual space of the book results in a process of incompleteness and deferment, and in a variable and open conformation between writing space and book space.

---

<sup>10</sup> One of the difficulties alluded to by Pessoa in the quoted passage derives from the fact that in the first stage of composition (1913–1920), he assigned the *Livro* to Vicente Guedes, whose style and psychology are significantly different from the book's later heteronym, Bernardo Soares, responsible for the second stage of composition (1928–1935). Vicente Guedes authored most of the "large texts" in the *Livro*. Editors have tried to solve the dilemma arising from those compositional differences in four different ways: Coelho (Pessoa 1982) arranges the texts according to a combination of chronology and thematic affinities, assigning the whole book to Bernardo Soares; Sobral Cunha (Pessoa 2013a) divides the texts into two groups and assigns each of them to Vicente Guedes and Bernardo Soares, while defining an internal textual order for each part based on thematic proximity; Zenith (Pessoa 2012) makes Bernardo Soares's pieces the structural axis of the book and relegates all the earlier large texts by Vicente Guedes to a final section in his edition; finally, after determining a likely date for all undated fragments, Pizarro (Pessoa 2013b) decides to follow a strictly chronological and genetic sequence, assigning all texts to Fernando Pessoa.

Thus, Pessoa's heteronymic split is not only the result of a retroactive effect of subjectivity produced by a given writing mode. It results also from the non-coincidence between the order of writing and the order of the book, which unfolds in an authorial self-consciousness as a product of the rules of writing and in an authorial self-consciousness as a product of the rules of the book. "Bernardo Soares" appears as a psychological entity that manifests itself in a given style and as the name of the potential author of a book, an author who is recursively a product of the book that he wants to produce, and that not only by the psychography of his writing mode. The implication is that the writing rules that define him as a heteronymic author are also a device for the production of a bibliographical coherence through which the heteronym edits himself as an author, thus determining the texts that are part of his book in

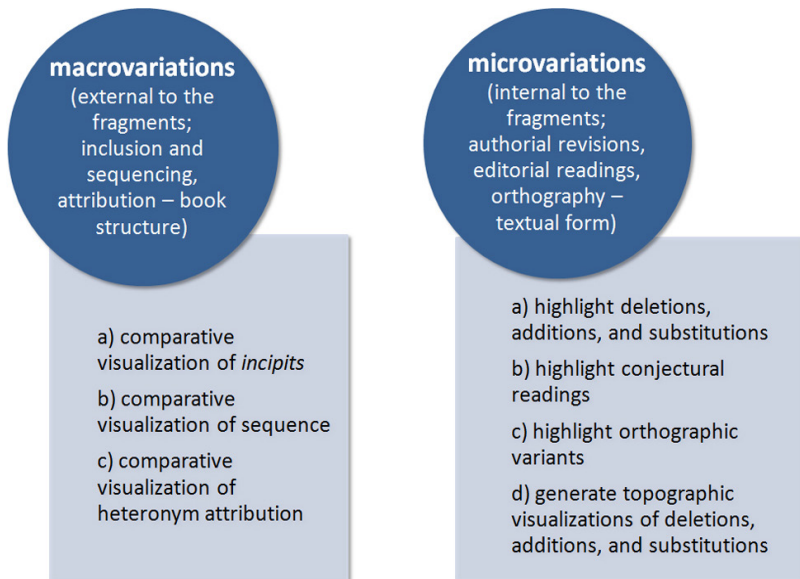


Figure 3: Authorial and editorial texts are encoded as a network of variations, both at the micro-scale of the scripting acts that constitute each fragment and at the macro-scale of codex sequence. Through algorithmic processing, the prototype interface gives readers the ability to visualize authorial and editorial variations in terms of both textual form and book structure.

```

138 <div xml:id="Fr311.TEXT">
139 <p>
140 Todos os dias acontecem no mundo<lb ed="#Fr311.WIT.MS.Fr311a.378"/>
141 coisas que não são
142 <app type="orthographic">
143 <rdg wit="#Fr311.WIT.MS.Fr311a.378 #Fr311.WIT.ED.CRIT.C">explicaveis</rdg>
144 <rdg wit="#Fr311.WIT.ED.CRIT.Z #Fr311.WIT.ED.CRIT.SC #Fr311.WIT.ED.CRIT.P">xeplicáveis</rdg>
145 </app> pelas<lb ed="#Fr311.WIT.MS.Fr311a.378"/>
146 leis que conhecemos
147 <app>
148 <rdg wit="#Fr311.WIT.MS.Fr311a.378 #Fr311.WIT.ED.CRIT.C #Fr311.WIT.ED.CRIT.Z">das</rdg>
149 <rdg wit="#Fr311.WIT.ED.CRIT.P #Fr311.WIT.ED.CRIT.SC">dos</rdg>
150 </app>
151 <app>
152 <rdg wit="#Fr311.WIT.MS.Fr311a.378">coisas<add place="above">seres</add></rdg>
153 <rdg wit="#Fr311.WIT.ED.CRIT.P #Fr311.WIT.ED.CRIT.SC">seres</rdg>
154 <rdg wit="#Fr311.WIT.ED.CRIT.Z #Fr311.WIT.ED.CRIT.C">coisas</rdg>
155 </app>. Todos<lb ed="#Fr311.WIT.MS.Fr311a.378"/>
156 os dias,
157 <app type="orthographic">
158 <rdg wit="#Fr311.WIT.MS.Fr311a.378 #Fr311.WIT.ED.CRIT.P #Fr311.WIT.ED.CRIT.C">falladas</rdg>
159 <rdg wit="#Fr311.WIT.ED.CRIT.Z #Fr311.WIT.ED.CRIT.SC">faladas</rdg>
160 </app> nos momentos, esquecem,<lb ed="#Fr311.WIT.MS.Fr311a.378"/>
161 e o mesmo
162 <app type="orthographic">
163 <rdg wit="#Fr311.WIT.MS.Fr311a.378 #Fr311.WIT.ED.CRIT.P">mysterio</rdg>
164 <rdg wit="#Fr311.WIT.ED.CRIT.C">misterio</rdg>
165 <rdg wit="#Fr311.WIT.ED.CRIT.Z #Fr311.WIT.ED.CRIT.SC">mistério</rdg>
166 </app> que as trouxe <lb ed="#Fr311.WIT.MS.Fr311a.378"/> as leva,
167 convertendo-se o segredo<lb ed="#Fr311.WIT.MS.Fr311a.378"/>
168 em esquecimento. Tal é a lei<lb ed="#Fr311.WIT.MS.Fr311a.378"/>

```

Figure 4: Excerpt from the XML-TEI encoding of the fragment ‘Todos os dias acontecem no mundo’ (BNP/E3, 5–23r): authorial and editorial textual versions are treated as variants and variations for encoding purposes. The <rdg> TEI element stands for reading and is used to represent both authorial and editorial micro variations. The editions and authorial sources are referred through the “wit” attribute. Additionally, a structured hierarchical nomenclature is used to identify witnesses, for instance, #Fr311.WIT.MS.Fr311a.378 denotes an authorial source (MS) witness (WIT) identified by Fr311a.378 of fragment 311 (Fr311), where 311 is an arbitrary number that identifies a particular XML file within the *LdoD Archive* system.

progress and, through this joint production of writing and codex, the subject of writing produces its self as the author of the book.

### Encoding and Visualizing Variation

As mentioned above, representation of the dynamics of variation in the *LdoD Archive* involves the consideration of two distinct levels. One is the level of micro-variations, i.e. variations that are internal to the fragments, such as authorial revisions, editorial readings of particular passages or orthographic variants that resulted from

reforms in spelling conventions.<sup>11</sup> The other is the level of macro-variations, i.e. variations that are external to the fragments, such as inclusion and sequencing of fragments, as well as heteronym attribution. In other words, the first type of variations results in a given textual form for each fragment or piece of writing, while the second type results in a given book structure for the entire corpus (Figure 3). How are these micro- and macro-variations represented in the *LdoD Archive*? How are authorial revisions and editorial variants marked and visualized in ways that enable readers to understand the writing and editorial processes at the scale of both textual form and book structure?

We may say that the representation of the genetic dimension takes place in the context of the work's socialized dimension, while the work's editorial forms can be perceived in the context of its genetic history. Revision processes in the autograph materials as well as variants and variations in editorial readings are encoded in the same XML file in a way that allows for both a single view of each autograph or editorial witness and comparative views of multiple witnesses (Figure 4). At the level of the header each TEI file contains the metatextual information required for comparing bibliographical features, such as "L. do D." markers in Pessoa's papers, numerical sequence of fragments in each edition, date of composition, or other metatextual attributes.

In our prototype of the *LdoD Archive*, visualization of variations takes place at the general level of the graphical user interface and within textual transcriptions. The user is allowed to move within each (authorial or editorial) textual witness and across different textual witnesses. This navigational strategy allows readers to see revision sites within authorial witnesses but also to generate

---

<sup>11</sup> A reformed orthography was introduced in Portugal in 1911, but Pessoa continued to write according to the earlier spelling conventions. In their editions of the *Livro*, Coelho (Pessoa 1982) and Pizarro (Pessoa 2010) follow Pessoa's orthography, while Sobral Cunha (Pessoa 2008) and Zenith (Pessoa 2012) have modernized Pessoa's spelling according to contemporary Portuguese orthography, i.e. the spelling agreement of the 1970s (the convention that has been used until now). We are currently (2010–2015) in a period of transition to a newly reformed orthography of the Portuguese language, which means that further spelling variations will be added to forthcoming editions of the *Livro*. In the *LdoD Archive* all spelling variations in both authorial sources and editorial transcriptions have been marked up.

The screenshot displays the LdoD Archive interface. At the top, there is a navigation bar with 'LdoD Archive' on the left and 'Logout' on the right. Below this, a menu contains options: 'About', 'Documents', 'Editions', 'Search', 'Virtual', and 'Admin'. A secondary menu includes 'Highlight Differences', 'Show Deleted', 'Highlight Inserted', 'Highlight Substitutions', 'Show Notes', and 'Facsimile'. The main content area is titled 'Todos os dias acontecem' and features a topographic transcription of an authorial witness on the left, overlaid on a digital facsimile of an autograph manuscript. The transcription shows text with various revision sites marked. To the right of the transcription is a text box containing the digital facsimile of the autograph manuscript. On the far right, there is a 'Virtual Editions' sidebar. It lists 'Authorial Sources' with 'BNP/E3, 5-23r' checked. Below that, it lists 'Expert Editions' with four entries: 'Jacinto do Prado Coelho' (with navigation arrows), 'Teresa Sobral Cunha' (with navigation arrows), 'Richard Zenith' (with navigation arrows), and 'Jérónimo Pizarro' (with navigation arrows). At the bottom of the sidebar, there are 'Virtual Editions' and 'EdTax' options, with an 'Add' button.

Figure 5: Screenshot of the prototype interface: a topographic transcription of authorial witness is shown against the corresponding digital facsimile of autograph manuscript (BNP/E3, 5–23r). Revision sites are XML encoded: deletions, additions and substitutions can be shown in the transcription. These micro-variations in each textual fragment are also mapped onto the macro-variations of codex structure in each edition: the right hand-side menu allows readers to navigate to the corresponding page in four different “Expert Editions”: Pessoa 1982, 2012, 2008, 2010. This menu further identifies the “Authorial Sources”, and also those “Virtual Editions” that the users of the archive have decided to make public.

comparisons between any 2, 3, 4, or N witnesses, when  $N > 5$  means that there are several authorial sources for the same fragments, comparing any given authorial witness against its editorial versions, and also editorial witnesses amongst themselves. This ability of examining the micro-variations in the textual form of each fragment across the database of witnesses is further contextualized, at the level of macro-variations, by the possibility of navigating within the bibliographical sequence offered by each scholarly edition (Figure 5). Buttons for showing revision sites (deletions, additions, and substitutions) and buttons for comparing transcriptions against the digital facsimiles of authorial witnesses allow users to move across all layers of variation from within a single screen. The right-hand menu provides immediate visualization of the relative position of each fragment within any given expert book edition of

the *Livro*, while the bottom of the page note provides other metatextual information concerning heteronym attribution or “*L. do D.*” mark (Figure 6). One-to-one or one-to-many comparisons between the archive’s transcription and the four editions are also supported. This principle also applies to each and all expert editions.

## Generating Variation

Our scholarly digital representation of authorial sources and expert editions as described in section 4 will provide the database materials for user-generated variations. Those user-generated variations on the *Livro* can result from either the editor-function or the author-function, both of which are explicit roles programmed in our virtual model of *Livro*. In this section we briefly explain the virtual editing and the virtual authoring modes, and the ways in which those new variations on *Livro* are visualized in the prototype interface.

Virtual editing results in the creation of a persistent virtual edition (Figure 7). A group of users can create and work together in a virtual edition; they are the virtual edition members. Virtual editions can be either private or public, where private virtual

The screenshot displays the LdoD Archive interface. At the top, there is a navigation bar with 'LdoD Archive', 'About', 'Documents', 'Editions', 'Search', 'Virtual', 'Admin', 'Logout', and a user profile icon. Below the navigation bar, there are two main columns. The left column shows the archive's transcription of the text 'Todos os dias acontecem' with highlighted variations in blue. The right column shows the first published edition of the same text. Below the text, there are two boxes containing identification and title information. On the right side, there is a sidebar with 'Authorial Sources' (BNP/E3, 5-23r), 'Expert Editions' (Jacinto do Prado Coelho, Teresa Sobral Cunha, Jerónimo Pizarro), and 'Virtual Editions' (EdTax).

Line-by-line  Align Spaces

Todos os dias acontecem

Todos os dias acontecem no mundo coisas que não são explicáveis pelas leis que conhecemos das coisas seres. Todos os dias, falladas nos momentos, esquecem, e o mesmo misterio que as trouxe as leva, convertendo-se o segredo em esquecimento. Tal é a lei do que tem que ser esquecido porque não pode ser explicado. A luz do sol continua regular o mundo visível visível. O alheio espreita-nos da sombra.

Todos os dias acontecem no mundo coisas que não são explicáveis pelas leis que conhecemos das coisas. Todos os dias, falladas nos momentos, esquecem, e o mesmo misterio que as trouxe as leva, convertendo-se o segredo em esquecimento. Tal é a lei do que tem que ser esquecido porque não pode ser explicado. A luz do sol continua regular o mundo visível. O alheio espreita-nos da sombra.

Identificação: bn-acpc-e-e3-5-1-85\_0045\_23\_t24-C-R0150  
Formato: Folha  
Material: Papel  
Colunas: 1  
LdoD: true  
Notas: Testemunho manuscrito a lápis., LdoD, Testemunho no recto de uma folha solta.

Título: Todos os dias acontecem  
Heterónimo: Bernardo Soares  
Número: 288  
Volume: II  
Página: 13  
Notas: [5-23. ms.];

Authorial Sources  
 BNP/E3, 5-23r

Expert Editions  
Jacinto do Prado Coelho  
 288  
Teresa Sobral Cunha  
 317  
Richard Zenith  
 424  
Jerónimo Pizarro  
 514

Virtual Editions  
EdTax

Figure 6: Screenshot of the prototype interface: comparison between the archive’s transcription of authorial witness and its corresponding form in the first published edition (Pessoa 1982). Textual variation sites are highlighted, which means the archive’s transcription is in itself a new editorial representation.





Figure 7: Screenshot of the prototype interface: the fragment is the number 1 of the EdTax virtual edition, and uses the Richard Zenith transcription where it is fragment number 424. A user, who is a member of the virtual edition, is currently adding an annotation to a part of the fragment. She can, eventually, also add a tag.

editions can only be visualized by their members. A virtual edition contains a set of fragments which are selected and ordered by its members. Virtual edition fragments use the archive's topographic transcription and/or the expert's editorial transcription and they can be enriched by annotations made by virtual edition members. An annotation can contain a comment and one or more tags, and it is associated to a part of the fragment transcription, the quote. The system also allows a virtual edition fragment to use annotations from another virtual edition fragment. The use of annotations and tags preserves the authorship of the original contributions. Therefore, when a virtual edition is built on top of another virtual edition, the source tags and annotations are inherited but cannot be changed in the context of the new virtual edition. This means that an inherited tag can be used in the new virtual edition to categorize a fragment or a part of a fragment but if it is changed or deleted in the source edition the change is reflected in the new virtual edition by, respectively, changing it or deleting all its uses. Obviously, inherited tags and annotations cannot be changed or deleted in the context of the new virtual edition though it is possible to create new tags and annotations.

In the context of a virtual edition, any of its members can associate annotations to parts of the fragments transcriptions. When visualizing a fragment of a virtual edition all the tags and contributors

The screenshot displays a web interface for comparing virtual editions. At the top, a navigation bar includes 'Arquivo LdoD', 'Acerca', 'Documentos', 'Edições', 'Pesquisa', 'Virtual', 'Admin', and 'Sair'. The main content area is titled 'Comparação de fragmentos de edições virtuais' and is divided into two columns: 'EdTax' and 'YetAnotherEd'. Each column contains a text fragment, a citation, a comment, the user who made the change, and tags. The 'EdTax' column shows a fragment about the sun and world, with a comment 'Existente uma...' and tags 'ars' and 'mundo'. The 'YetAnotherEd' column shows a fragment about things happening in the world, with a comment 'sobre a visibilidade...' and tags 'mp' and 'mundo'. On the right side, there is a sidebar with 'Fontes Autorais' (BNP/E3, 5-23r), 'Edições dos Peritos' (Jacinto do Prado Coelho, Teresa Sobral Cunha, Richard Zenith, Jerónimo Pizarro), and 'Edições Virtuais' (EdTax, YetAnotherEd).

Figure 8: Screenshot of the prototype interface: comparing a fragment in two virtual editions, YetAnotherEd and EdTax, their quotes, comments, tags and contributors.

are highlighted. Since a fragment can be annotated in the context of different virtual editions it is possible to compare these contributions, as shown in Figure 8.

The author-function provides another role within the archive for generating variations. Variation is understood here as a mode of literary production that is self-conscious about its intertextual dimension, and takes fragments or passages from the *Livro de Desassossego* for further textual production. Since the author-function is not implemented in the prototype yet, we will simply list a series of different techniques for writing variations on *Livro* fragments contemplated by our virtual model. The virtual writing functions should satisfy the following conditions: selecting sources (selecting transcriptions according to a particular edition); defining source elements at different levels of granularity (from single word to phrase to sentence to entire fragment); defining various degrees of human-machine collaboration (from blog-like pieces of human-authored text to entirely computer-authored permutations); defining anchors and links; defining order; and defining heteronym attribution.

User-generated editions and user-authored texts will become part of the virtual layer of the *LdoD Archive*, available for further reading and manipulation. The fully dynamic nature of *LdoD Archive* as a Web 2.0 experimental project results from this particular integration of the scholarly encoding of the work's authorial

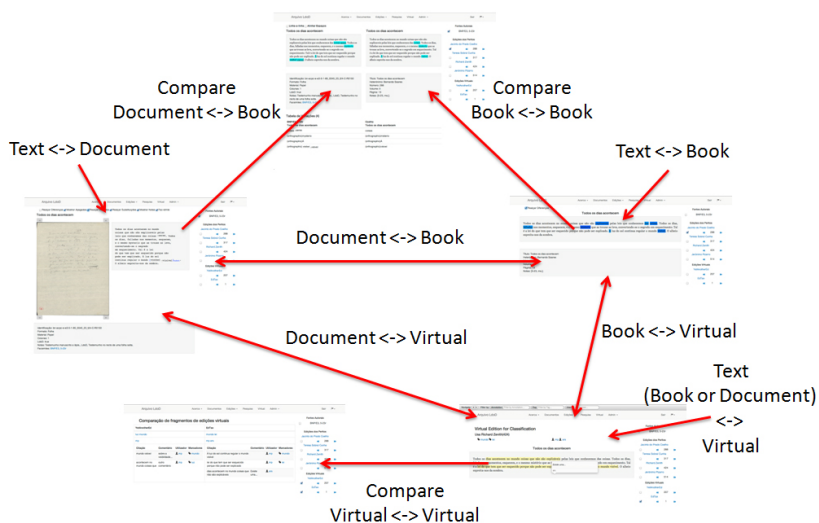


Figure 9: Relations at the micro- and macro-variation levels.

and editorial archive with the algorithmic and collaborative simulation of the literary processes of reading, editing, and writing.

## Discussion

For a fuller understanding of the potentialities of this approach we illustrate in Figure 9 how the different graphical interfaces map the interrelations among representations of Document, Text, and Book, at the level of both micro- and macro-variations. Through these interfaces it is possible to visualize Text $\leftrightarrow$ Document relations when the facsimile is seen side-by-side with its topographic transcription. From this visualization it is possible to navigate to an interface where the Document's textual transcription is compared to the Book's expert editorial transcriptions. Each expert editorial transcription is situated within a bibliographical sequence in its respective edition. It is also possible to navigate to Book and Virtual interfaces for the same fragment.

When looking at the macro-variation level it is possible to visualize different aggregations of fragments at several levels, including Book, Document, Virtual, Contributor, and Classification. These visualizations enable large-scale comparisons across tables of contents and lists of texts that will give readers a sense of

bibliographical structure, textual sequence and semantic classifications of various versions of the *Book of Disquiet*, in both expert and virtual editions.

Micro-variations across textual transcriptions and macro-variations across bibliographical structures in the work's genetic and editorial archive are displayed through a network of shifting perspectives. This network of shifting perspectives allow users of the archive to see Pessoa's writing process and his changing and variable plans for the *Livro*. At the same time, readers will become aware of the conjectural nature of the editorial solutions of the expert critical editions for producing a structured textual and bibliographical form out of a half-finished and fragmentary work. Further micro- and macro-variations will result from the archive's socialization of editorial and authorial acts of production at the virtual level, opening up the work's existing archive to future appropriations and transformations.

### Conclusion

Our attempt at creating a constellated archive of authorial and editorial witnesses follows from our rationale for integrating the work's authorial and editorial history. It is also motivated by a model for virtualizing the *Livro de Desassossego* in ways that take advantage of the collaborative affordances of the medium to explore both the nature of writing and the nature of the book. Visualization of variation in the *LdoD Archive* takes places at three related levels: first, authorial witnesses are represented as digital image facsimiles (manuscripts, typescripts, printed states), according to a documentary editing approach; secondly, authorial witnesses and expert editions are transcribed and TEI encoded using XML tags for representing textual features, including revision sites and variants; thirdly, the user interface is designed for enabling readers to navigate between micro and macro-level variations in ways that enhance their understanding of both the work's writing process and its bibliographical structures in various editorial instantiations.

Generation of further variation in the *LdoD Archive* implies that readers are given the possibility of performing bibliographical and textual interventions according to editor-function and author-function. This open use of authorial and editorial witnesses contained

in the scholarly level of the archive takes place at the virtual level, where readers will be able to select, order, annotate and comment on particular fragments. They can also transition from an editorial to an authorial role and write variations and extensions based on particular passages or fragments. Editing and writing take place at this virtual dimension as actual practice and simulation of literary processes. The work's genetic and editorial history is recontextualized in the cooperative networked textual environment maintained by our database and algorithmic model of literary performance. Through XML encoding and programmability the digital environment is experimentally designed for a critical engagement with textual variation, transmission and production in ways that allow us to experiment with writing, editing and meta-editing beyond the bibliographical horizon.

### Bibliography

- Bénel, Aurélien and Christophe Lejeune. 2009. "Humanities 2.0: Documents, Interpretation and Intersubjectivity in the Digital Age". *International Journal on Web Based Communities*, 5, pp. 562–76.
- Bryant, John. 2012. *The Fluid Text: A Theory of Revision and Editing for Book and Screen*. Ann Arbor: The University of Michigan Press.
- Fraistat, Neil and Steven Jones. 2009. "Editing Environments: The Architecture of Electronic Texts". *Literary and Linguistic Computing*, 24(1), pp. 9–18.
- Pessoa, Fernando. 1982. *Livro do Desassossego*. 2 vols. Ed. Jacinto do Prado Coelho. Lisboa: Ática.
- . 2012. *Livro do Desassossego*. Ed. Richard Zenith. Lisboa: Assírio & Alvim.
- . 2008. *Livro do Desassossego*. Ed. Teresa Sobral Cunha. Lisboa: Relógio d'Água.
- . 2010. *Livro do Desassossego*. 2 vols. Ed. Jerónimo Pizarro. Lisboa: Imprensa Nacional-Casa da Moeda.
- . 2013a. *Livro do Desassossego*. Ed. Teresa Sobral Cunha. Lisboa: Relógio d'Água.
- . 2013b. *Livro do Desassossego*. Ed. Jerónimo Pizarro. Lisboa: Tinta da China.

- Pizarro, Jerónimo. 2012. *La Mediación Editorial: Sobre la Vida Póstuma de lo Escrito*. Madrid: Iberoamericana / Vervuert.
- . 2013. “Os Muitos Desassossegos”. In *Comunicações do III Congresso Internacional Fernando Pessoa*, Lisboa: Casa Fernando Pessoa, pp. 1–15, <[http://www.congressointernacionalfernandopessoa.com/comunicacoes/jeronimo\\_pizarro.pdf](http://www.congressointernacionalfernandopessoa.com/comunicacoes/jeronimo_pizarro.pdf)>.
- Portela, Manuel. 2013. “Nenhum Problema Tem Solução: Um Arquivo Digital do *Livro do Desassossego*”. *MatLit*, 1(1), pp. 9–33, <<http://iduc.uc.pt/index.php/matlit/article/view/1618>>.
- Portela, Manuel and António Rito Silva. 2014. “A Model for a Virtual LdoD”. *Digital Scholarship in the Humanities*, 30, pp. 354–70.
- Robinson, Peter. 2013. “Towards a Theory of Digital Editions”. *Variants*, 10, pp. 105–131.
- Sepúlveda, Pedro. 2013. “Listas do Desassossego”. *MatLit*, 1(1): 35–55, <<http://iduc.uc.pt/index.php/matlit/article/view/1616>>.
- Shillingsburg, Peter L. 2006. *From Gutenberg to Google: Electronic Representations of Literary Texts*. Cambridge: Cambridge University Press.
- Siemens, Ray, Meagan Timney, Cara Leitch, Corina Koolen, and Alex Garnett, with the ETCL, INKE, and PKP Research Group. 2012. “Toward Modeling the Social Edition: An Approach to Understanding the Electronic Scholarly Edition in the Context of New and Emerging Social Media”. *Literary and Linguistic Computing*, 27, pp. 445–61.
- Silva, António Rito and Manuel Portela. 2013. “Social Edition 4 *The Book of Disquiet*: The Disquiet of Experts with Common Users”. In Matthias Korn, Tommaso Colombino, and Myriam Lewkowicz (eds.), *ECSCW 2013: Adjunct Proceedings The 13th European Conference on Computer-Supported Cooperative Work*. Department of Computer Science, Aarhus University, pp. 45–50, <[http://cs.au.dk/~mkorn/ECSCW\\_2013\\_Adjunct\\_Proceedings-web.pdf](http://cs.au.dk/~mkorn/ECSCW_2013_Adjunct_Proceedings-web.pdf)>.
- . 2015. “TEI4LdoD: Textual Encoding and Social Editing in Web 2.0 Environments”, *Journal of the Text Encoding Initiative*, 8, <<http://jtei.revues.org/1171>>.
- Silvestre, Osvaldo Manuel. 2014. “O Que Nos Ensinam os Novos Meios sobre o Livro no *Livro do Desassossego*”. *MatLit*, 2 (1), pp. 79–98, <<http://iduc.uc.pt/index.php/matlit/article/view/1818>>.
- Vanhoutte, Edward. 2006. “Prose Fiction and Modern Manuscripts: Limitations and Possibilities of Text-Encoding for Electronic

- Editions". In Lou Burnard, Katherine O'Brien O'Keeffe and John Unsworth (eds.), *Electronic Textual Editing*. New York: Modern Language Association of America, pp. 161–80.
- Werner, Marta L. 2011. "'Reportless Places': Facing the Modern Manuscript". *Textual Cultures*, 6 (2), pp. 60–83.
- Zenith, Richard. 2013. "*Livro do Desassossego*: o romance possível (var.: impossível)". In *Comunicações do III Congresso Internacional Fernando Pessoa*, Lisboa: Casa Fernando Pessoa, pp. 1–12, <[http://www.congressointernacionalfernandopessoa.com/comunicacoes/richard\\_zenith.pdf](http://www.congressointernacionalfernandopessoa.com/comunicacoes/richard_zenith.pdf)>.