'Status Quaestionis' of Dantean Digital Resources. Some Corrective Theoretical Perspectives and a Case Study (Database Allegorico Dantesco)

Matteo Maselli
University of Macerata, mattmas91@yahoo.it

Recommended Citation

'STATUS QUAESTIONIS' OF DANTEAN DIGITAL RESOURCES. SOME CORRECTIVE THEORETICAL PERSPECTIVES AND A CASE STUDY (DATABASE ALLEGORICO DANTESCO)

MATTEO MASELLI, University of Macerata[1]

This paper reflects on the current configuration of Dantean digital resources and proposes some possible perspectives to implement their functionality (disambiguation, RDF and Semantic Web, Distant & Close Reading, attention to the processes of digitization of texts). The second part of the essay explains the structure and presents the basic functionality of the *Database Allegorico Dantesco* (DAD) a new repository on Dantean allegory.

Keywords: Digital Humanities, Semantic Web, Dantean Tools, Distant & Close Reading, Allegory

1. *Introduction*

Those who in recent years have had the opportunity to consult the ever-growing bibliography devoted to the relationship between Dante and the Digital Humanities will certainly have noticed some recurring aspects. I am referring to the tendency to critique what is available online and to insist that those tools have been programmed only for the utility of users and not to replace traditionalist practices of study. They claim this despite the fact that, as Akash Kumar has well pointed out recently, "[d]igital projects that center on Dante do not, by their very definition, attempt to subvert the canon."[2] By this I do not want to diminish the tendency to offer

---

[1] The author sincerely thanks Elizabeth Coggeshall and Akash Kumar, important interlocutors throughout the writing of this work. Thanks go especially to Elizabeth Coggeshall, always ready to dispense wise suggestions, who enabled me to structure the text as it now stands.

[2] Akash Kumar, "Digital Dante," in *The Oxford Handbook of Dante*, ed. M. Gragnolati, E. Lombardi, F. Southerden (Oxford: Oxford University Press, 2021), 97; see also Akash Kumar, Julie Van Peteghem, "Digital Dante. Reimagining Dante in a

~ 206 ~

an up-to-date sitography on Dante, even more so when contextualized at the historiographical level.[3] I am aware that if these sitographies continue to be composed it is because while we pay attention to Dante's digital resources, we often ignore their actual functionality and thus potential for study. This is certainly no small deficiency, since exploiting the possibilities of the Web represents an obvious advantage over the older generations of Dante scholars who, as Michele Barbi wrote in *Nuovi problemi della critica dantesca* (1932), were "uomini di grande ingegno e di grande dottrina, [ma] quasi tutti autodidatti […] [che lavorarono] con mezzi scarsissimi e fra mille difficoltà."[4]

What must be emphasized, however, is that we have now reached a point in the development of international *dantistica* where a merely descriptive approach to the tools of consultation and interrogation of Alighieri's works seems no longer sufficient. In fact, studies on Dantean sitographies often limit themselves to cataloging the available resources and pointing out their basic functions, but without questioning their limitations and possible solutions to modernize and make them more receptive to users' needs.

Virtual World," *L'Illustrazione. Rivista del libro a stampa illustrato* 5 (2021): 211-222.

[3] Among the various works available, I would highlight the following: Simona Brambilla, "Dante in internet," in *Atti del Convegno Internazionale "Commentare Dante oggi,"* ed. J. Kelemen, J. Nagy (Budapest: Eötvös University Press, 2015), 151-157; Riccardo Castellana, "Risorse digitali dantesche. Testi, commenti, filologia," *Allegoria. Per uno studio materialistico della letteratura* 48 (2004): 96-124; Elizabeth Coggeshall, "Dante oggi: l'Inferno diffondibile," *Italianistica. Rivista di letteratura italiana* 49 (2020): 73-88; Crystal Hall, "Digital humanities and Italian studies. Intersections and oppositions," *Italian Culture* 37 (2019): 97-115; Giuseppe Marrani, "Dante nel web," in *Dante nelle scuole. Atti del Convegno di Siena, 8-10 marzo 2007*, ed. N. Tonelli, A. Milani (Florence: Franco Cesati Editore, 2009), 179-196; Matteo Maselli, "Per una rassegna degli strumenti della critica dantesca. Dai repertori testuali ai dispositivi digitali," *Paratesto* 18 (2021): 299-337; Florinda Nardi, "Dante online. 'Per l'alto mare aperto.' Dante naviga sul World Wide Web," *Dante. Rivista Internazionale di Studi su Dante Alighieri* 1 (2004): 143-155; Teresa Nocita, "Da Internet. Siti monografici per gli studi di letteratura medievale. 'Auctoers' volgari. Dante," *Critica del testo* 3 (1999): 1233-1239; Mattia Slavemini, "Un viaggio nel viaggio. Dante in rete oggi," *Dante. Rivista Internazionale di Studi su Dante Alighieri* 13 (2016): 153-159; Gaia Tomazzoli, "Digital resources for Dante Studies. A critical survey," *Digital Philology. A Journal of Medieval Cultures* (forthcoming); Paolo Zara, "Sitografia su Dante," *Lineatempo. Itinerari di ricerca storica e letteraria* 1 (2004): 48-49.

[4] Michele Barbi, "Nuovi problemi della critica dantesca," *Studi Danteschi* 16 (1932): 37; Still valid is Jerome McGann's observation that humanities research will not take the use of digital technologies seriously until it can be demonstrated how they can improve the processes of exploring and interpreting texts [see Jerome McGann, "Preface," in *Radiant Textuality. Literary Studies after the World Wide Web* (New York: Palgrave MacMillan, 2004), XII].

~ 207 ~

Instead, we need a critical approach, in the sense of analyzing the possibilities for the development of what universities and research centers have so far programmed and released on the Web. Therefore, the present paper sets out to do just that, namely, not to represent the now overused list of Dantean digital resources, but to put forward some suggestions that could improve them.

Before proceeding it is necessary to make an important explanation about the type of investigation that will be proposed in the following pages and that goes to better define the focus of this paper. It is worth pointing out that the suggestions that will be put forward to better manage possible structural deficiencies in some of Dantean digital resources will concern the structural logic of them alone. This means that if methodologies of digital investigation will be mentioned, it will be done without the claim of examining them in the light of the most recent critical bibliography, but they will be taken into consideration limitedly to those aspects that are more functional to better highlight the unexpressed potentialities of the evaluated resources.

The aim of this paper is neither to trace a history of the contribution of Digital Humanities to international *dantistica* nor to show how the exam of Dante's works might benefit from the use of specific computer methodologies.[5] Such examinations, certainly worthy of consideration, could easily be conducted without placing Dantean applications at the center of the inquiry. That is, to ask why there is a reluctance to study Dante by resorting, for example, to computational analysis does not necessarily entail an awareness of how well a Dantean site meets the most recent *Guidelines for Electronic Scholarly Edition*. Therefore, aware that what will be said certainly cannot be taken as an all-encompassing solution to the various problems found in Dantean digital resources, it is hoped that the observations offered will at least have the merit of showcasing some possibilities for intervention to counteract the physiological obsolescence of digital products.

Therefore, I will first propose a schematic for evaluating the digital tools related to Dante's works. I will then examine some of

---

[5] In this regard, there is no shortage of recent proposals to apply new resources to the study of Dante's works. In addition to what will be said about the *Database Allegorico Dantesco* (see section 3), consider the following recently published works: Paola Italia, Sara Obbiso, Roberta Priore, "Analisi stratigrafiche e 3D. Casi di studio da Dante a Leopardi," *Umanistica Digitale* 12 (2022): 65-86; Ugo Conti, "Combining rhythmic and syntactic analysis: an experiment on Dante's Comedy with the new tool TRIARS," *Umanistica Digitale* 13 (2022): 49-68.

~ 208 ~

available strategies (like RDF and Semantic Web ontologies) that are currently underutilized in the study of Dante's works, but which could be profitable to digital scholars interested in researching the poem and its legacy. Furthermore, I will offer a discussion of a digital work-in-progress, the *Database Allegorico Dantesco*, which I have designed to meet the needs of those interested in the study of the *Commedia*'s allegorism.

2. *Evaluating Dantean sites*

A first evaluative approach may take into account the basic logic promoted by the MLA (Modern Language Association) with the release of the *Preliminary Guidelines for Electronic Scholarly Editions*. Even if these are addressed to the ecdotic modes of digital publishing practice, the principles outlined in the 2022 version[6] of the *Guidelines* can be applied broadly to more projects than critical digital editions. The principles in the 2022 *Guidelines* are: accuracy, adequacy, appropriateness, internal consistency, and explicitness.[7] These principles clearly articulate a set of parameters that can also be applied to the evaluation of Dantean sites. Even more pertinent, however, are those criteria that take into account user experience: the transparency of sources and bibliographic references (transparent); the constant site's updating (maintained); the availability of the site's resources across multiple languages (multilingual); the accessibility of resources (accessible); the site's responsiveness, which implies a small number of actions to get to the object of the search (responsive); interoperability (interoperable); clarity (managed); and content preservation and URL stability (preserved).

Considering these guiding criteria, it becomes quite evident that few of the major digital tools devoted to Dante and his work scrupulously comply with all the parameters illustrated in the *Guidelines* (e.g., the *Dante Lab Reader* or its predecessor *Dartmouth Dante Project*). Far more numerous, on the other hand, are the cases in which the resources are deficient in only one item (*DanteSources*, *Digital Dante*, *Bibliografia Dantesca Internazionale*, *Dante Online*, *Monarchia Digital Edition*, *Dante Today*, *Princeton Dante Project*) or in more than one structural feature

---

[6] See the following link: https://www.mla.org/Resources/Guidelines-and-Data/Reports-and-Professional-Guidelines/Publishing-and-Scholarship/Guidelines-for-Editors-of-Scholarly-Editions (last revised 4 May 2022).
[7] Tiziana Mancinelli, Elena Pierrazzo, *Che cos'è un'edizione scientifica digitale* (Rome: Carocci Editore, 2020), 100-101; Deborah Parker, "Guidelines for the evaluation of Digital Humanities Projects," *ADFL Bulletin* 41 (2009): 67-75.

~ 209 ~

(*Dante Medieval Archive*, *Dante Search*, *Illuminated Dante Project*, *Intertextual Dante*, Prue Shaw's Digital Edition of the *Commedia*, *The World of Dante*, *Vocabolario Dantesco*) (see fig. 1 below).

Dantean Tool	Transparent	Maintained	Multilingual	Accessible	Responsive	Interoperable	Managed	Preserved
Bibliografia Dantesca Internazionale	X	X	X	X	X		X	X
Commedia - Digital Edition (P. Shaw)	X	X		X			X	X
Dante Lab Reader	X	X	X	X	X	X	X	X
Dante Medieval Archive	X		X	X	X		X	X
Dante Online	X		X	X	X	X	X	X
Dante Today	X	X		X	X	X	X	X
DanteSearch	X	X		X			X	X
DanteSources	X		X	X	X		X	X
Dartmouth Dante Project	X	X	X	X	X	X	X	X
Digital Dante	X	X		X	X	X	X	X
Illuminated Dante Project	X	X	X	X			X	X
Intertextual Dante	X			X	X		X	X
Monarchia - Digital Edition	X	X	X	X	X		X	X
Princeton Dante Project	X		X	X	X	X	X	X
The World of Dante	X			X	X	X	X	X
Vocabolario Dantesco	X	X		X	X		X	X

Fig. 1 – Illustrative outline of the structural features of the main Dantean sites

Even in the case of project infrastructures that have already been certified as the final products of a concluded research cycle, such an assessment can nevertheless benefit the planning of future initiatives.[8] It is good to start order critical interventions by priority, namely those circumstances involving interaction with texts through user-set searches.

2.1. *Logical and structural expedients*

Projects that aim at querying text sequences (e.g. the *Dartmouth Project* and *Dante Medieval Archive*) can be assessed by the accuracy of the output of allowed search terms. It is therefore normal to expect an updated efficiency of the querying granted by the machine.[9] However, this is exactly the operational area in which it is possible to discern the greatest urgency for interventions aimed at remodeling the setting of digital applications. What most reduces the accuracy of search results concerns the possible ambiguity of search terms (background noise).[10]

---

[8] See Elizabeth Coggeshall's text in *Bibliotheca Dantesca*'s cluster on Dante and Digital Humanities (no. 5, 2022).

[9] Obsolescence as a result of a failure to update is another of the pivotal issues to be considered because a failure to revise in line with the times destines a digital resource for a very likely demise (this is, for example, the case, at least at the time this note is being written, with *Mapping Dante*) or an equally severe aging that leads to a near-uselessness of the tool (this is the case with Otfried Lieberknecht's *Dante Site*).

[10] See Michelangelo Zaccarello, *Ostacoli alla ricerca web: ambiguità e omonimia*, in *Leggere senza libri. Conoscere gli e-book di letteratura italiana* (Florence: Franco Cesati Editore, 2020), 81-83. Although the study that is set up by Zaccarello concerns research on the Web and not that accomplished in a controlled collection of texts such as a database might be, the principles described are equally applicable to the latter case.

~ 210 ~

Let us imagine, for example, that we have at our disposal an archive which houses complete digitization of the *Giuntina di rime antiche*, a collection of poetic texts by various authors dated to the first thirty years of the 1500s. Assume that we want to search it for all occurrences of the name "Dante," intending to index textual loci which refer to Alighieri alone. The software that will process the submitted input, contrary to the user's intentions, will also return passages in which Dante da Maiano is quoted, which is in fact part of the *Giuntina*'s sampling. Thus, the digital application would be unable to disambiguate the homonymy between the two authors, whose identities are automatically superimposed.

Let's assume now that we decide to modify our search, specifying "Dante Alighieri" instead of the ambiguous "Dante." But what if our author has sought to avoid redundancy by replacing "Dante" with the poet's other epithets, like "Sommo poeta," "autore fiorentino," or the entirely generic "poeta"? In this case, even a highly targeted search under the label "Dante Alighieri" will not return passages which nonetheless refer to him. Thus, it is clear how the risks of misleading or incomplete examination can seriously impact the quality of the output enabled by the applications available today.

Shortcomings such as these are essentially due to the text-based nature of the catalogued material. This means that the results of querying these texts are derived simply through the mediation of single or combined keywords. But the polysemous nature of Dante's works calls for more accurate search types, characterized by a reduced rate of ambiguity in the search results.[11] One possible solution points to the adoption of the principles of the Semantic Web and the grammar on which it depends.[12] The latter uses a data model—the Resource Description Framework (or RDF)—for the formal representation of knowledge, thus defining a syntax with which machine-readable messages (so-called "linked data") can be produced. What is most consequential in the use of linked data is the way it configures shared vocabularies ("ontologies") from

---

[11] As will be shown later with the *Database Allegorico Dantesco*, one possibility for allowing the search to be more accurate could be automatic suggestion by the resource of concepts not considered by the user.

[12] Tim Berners-Lee, James Handler, Ora Lassila, "The Semantic Web," *Scientific American* 5 (2001): 28-37.

~ 211 ~

which one may derive the linguistic entities that establish the syntactic relationships underlying the messages to be disseminated.[13]

Advocated by the World Wide Web Consortium (W3C), the RDF provides an abstract model for expressing a potentially infinite number of assertions about everything around us. Such assertions (statements) are structured as triples, indicating a binary relationship (predicate) between two entities (subject; object) contained in the reference ontology. Statements that employ the RDF allow to fix the natural language disambiguation: each element of the statement created with the Semantic Web ontology is associated with a unique recognition identifier (Internationalized Resource Identifiers or IRIs, otherwise known as Uniform Resource Identifiers or URIs). The sentence "Dante è l'autore della *Commedia*" can thus be identified as a structure consisting of three elements, to which as many IRI-URIs correspond: "Dante" (the subject) is in a relationship of authorship (predicate) with "the *Commedia*," the object of his writing (fig. 2).[14]
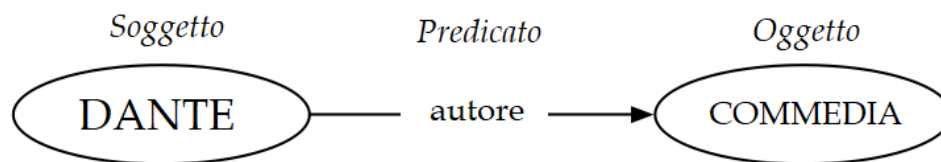

Fig. 2 – Example of triple (Subject + Predicate + Object)

---

[13] More specifically, ontology can be defined as the set of basic terms and relations that constitute the vocabulary of a specific area and the rules for combining terms and relations to extend the vocabulary [Robert Neches, "Enabling Technology for Knowledge Sharing and Reuse," *AI Magazine* 12 (1991): 36–56] or as a common and shared way of understanding a field that can be communicated between people and application systems [Thomas R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing," *IJHCS* 43 (1994): 907]; See also Michael Gruninger, Mike Ushold, "Ontologies: Principles, methods and applications," *Knowledge Engineering Review* 11 (1996): 93-136. On this topic see the Susan Hockey Lecture in Digital Humanities 2018 entitled "What can be said, can be said clearly? The role of ontologies in the Digital Humanities" given by Carlo Meghini and available at the following link: https://www.ucl.ac.uk/digital-humanities/events/SusanHockeyLecture/2018; Valentina Bartalesi, Carlo Meghini, "Using an ontology for representing the knowledge on literary texts: the Dante Alighieri case study," *Semantic Web* 1 (2015): 1–10.

[14] It should be noted, however, that RDF may have limitations in its use, which can be overcome by switching to OWL (Ontology Web Language), which allows for the harmonization and organization of assertions made in RDF without making multiple assertions about URIs representing individual concepts but as belonging to classes characterized by specific properties.

~ 212 ~

Should one decide to weave a series of statements written with the data model RDF, one would be defining an extensive knowledge map (network) in which data would be carefully collected and assembled into semantic categories on which the machine could then operate with targeted interventions fostering fundamental knowledge interoperability (fig. 3).[15]
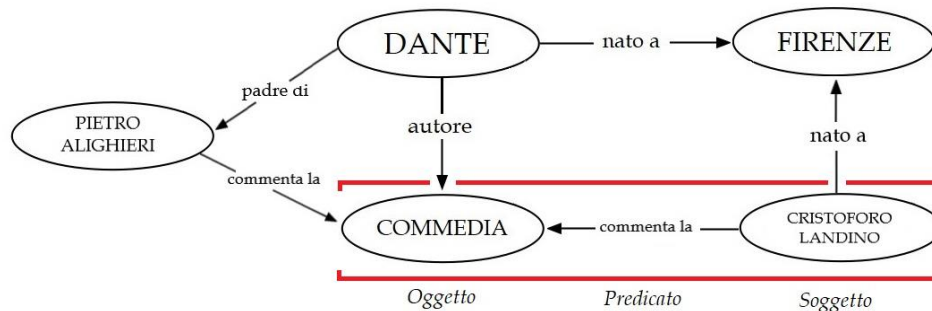

Fig. 3 – Example of networking triples in traced but expandable pathways that ensure relevance of research and interconnectedness of knowledge

By creating order in the scattered mass of online data, the RDF will thus be able to guide users along paths inherent to their research and enable easy interaction across multiple subject areas. However far such a scenario has not yet been reached, such a prospect could certainly benefit even those who are reluctant to relate a digital point of view to textual materiality.

Given the apparent advantages of RDF, one hopes that the same might soon become a common protocol in Dante studies,[16] although there are objective logistical difficulties, particularly that the use of RDF and Semantic Web ontologies presupposes, even more than in other circumstances, close cooperation between professional computer scientists and humanists. Despite these obvious challenges, there is already at least one major digital project that plans to employ Semantic Web ontologies to analyze Dante's works

---

[15] Valentina Bartalesi, Carlo Meghini, Daniele Metilli, "A Conceptualisation of Narratives and Its Expression in the CRM (Conceptual Reference Model)," *International Journal of Metadata, Semantics and Ontologies* 12 (2017): 35-46; *The Semantic Web. Research and Applications. 6th European Semantic Web Conference (Heraklion, May 31-June 4, 2009),* ed. L. Aroyo, E. Hyvönen (Berlin: Springer, 2009); Benjamin M. Schmidt, *'*Do Digital Humanists Need to Understand Algorithms?," *Debates in Digital Humanities* (2016): https://dhdebates.gc.cuny.edu/read/untitled/section/557c453b-4abb-48ce-8c38-a77e24d3f0bd##ch48.

[16] Marco Passarotti, Giulia Pedonese, Rachele Sprugnoli, "Le opere latine di Dante tra annotazione linguistica e web semantico," *Linguistica e Letteratura* 46 (2021): 45-71.

~ 213 ~

and their sources: this is the *Hypermedia Dante Network*, a collaboration between the University of Pisa, ISTI (Istituto di Scienza e Tecnologie dell'Informazione)–CNR (Consiglio Nazionale delle Ricerche) and the Universities of Bologna, Turin and Naples supervised by Michelangelo Zaccarello — which aims to expand the potential of *DanteSources*.[17]

Another highly suggestive possibility concerns how information is visualized and implicitly analyzed, which is not necessarily far from the semantic mapping just mentioned. By this I do not refer to the visual legacy of the *Commedia*, since nearly all Dante sites—with only a few notable exceptions, such as the *Princeton Dante Project*—propose proper placement of available material in corresponding graphical interface. Instead, what I am referring to here is the utility of distant reading, a practice which, when compared to close reading, permits new ways of interacting with cataloged data.[18]

Distant reading is a practice that offers a panoramic and comprehensive look at the objects of its investigation, a global view of the text complementary to the granular examination of close reading that instead permits the reader to highlight relationships between and within large data sets.[19] Distant reading allows the

---

[17] Paola Andriani, Valentina Bartalesi, Carlo Meghini, Mirko Tavoni, "Towards a semantic network of Dante's works," *Digital scholarship* 30 (2015): 28-35; Carlo Meghini, Mirko Tavoni, Michelangelo Zaccarello, "Mapping the Knowledge of Dante Commentaries in the Digital Context: A Web Ontology Approach," *Romanic Review* 1 (2021): 138-157; Valentina Bartalesi, Nicolò Pratelli, Carlo Meghini, Daniele Metilli, Gaia Tomazzoli, Leyla M.G. Livraghi, Michelangelo Zaccarello, "A formal representation of the Divine Comedy's primary sources. The Hypermedia Dante Network ontology," *Digital scholarship in the Humanities* 0 (2021): 1-14; see also Gaia Tomazzoli's piece "Intertextuality in Dante's Commedia: Hypermedia Dante Network" in "Notes" section of the *Bibliotheca Dantesca* (no. 5, 2022) essay cluster.

[18] For an especially compelling use case, see Jacob Blakesley's contribution to this cluster (*Bibliotheca Dantesca*, no. 5, 2022), on distant reading translations of Dante's *Commedia*.

[19] Franco Moretti, "Conjectures on World Literature," *New Left Review* 1 (2000). Also by Moretti, consider the following works: *A una certa distanza. Leggere i testi letterari nel nuovo millennio* (Rome: Carocci Editore, 2020); *Falso movimento. La svolta quantitativa nello studio della letteratura* (Milan: Nottetempo, 2022); William E. Underwood, "A Genealogy of Distant Reading," *Digital Humanities Quarterly* 11 (2017): http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html; see also the paper presented by Ginestra Ferraro *et al.* at The Alliance of Digital Humanities Organizations virtual Conference (22-24 July 2020) titled "The place of models and modelling in Digital Humanities: Intersections with a Research Software Engineering perspective": https://hcommons.org/deposits/item/hc:31819/.

~ 214 ~

researcher to extract and organize knowledge according to taxonomies, or to profile the language of a particular author, text, or group of texts by one or more authors. These relationships are visualized in clear graphical representations (e.g., maps, diagrams, tables, graphs).

To date, there are few Dantean resources that have attempted to leverage distant reading. One site, *DanteSources*, does so with bar graphs describing the presence of textual sources in Dante's works. Another, *The World of Dante*, offers an interesting timeline on the life of the Florentine poet and the main historical events contextual to it. On the other hand, a large percentage of Dantean applications share the advantages of close reading, a practice which is particularly suited to research that aims not only at a study of the deep content of the text but also its intertextual/intratextual dynamics. Indeed, close reading, in which the researcher's gaze is directed at the text's compositional units with absolute proximity, presupposes a parallel and meticulous knowledge of the secondary literature. Sites that enhance the latter, can then be placed in the reading category we are talking about (to name just a few we can think of the *Dartmouth Dante Project*, the *Dante Lab Reader*, the *Dante Medieval Archive* or the *Intertextual Dante*).

Aware of the advantages expressed by both close and distant reading—especially when employed in tandem—, one hopes that those in charge of the various networked Dantean projects will soon expand their methodological toolkit to include hybrid data management methodologies. Contrary to what one may believe, distant reading is not antithetical to close reading; rather, each complements the other by bridging their respective limitations.

I would proffer a final suggestion, primarily for future digitization projects. This concerns the problem that Paola Italia and Roberto Calasso have called the "disappearance of the paratext."[20] That is, the processes of text digitization entail the almost automatic loss of all those paratextual traces that, however integral they are to the analog reading experience, are not necessarily so for the digital one.[21] The seriousness of such a loss is quickly demonstrated if one

---

[20] Roberto Calasso, "Filologia editoriale. Dialogo con Paola Italia e Francisco Rico," *Ecdotica* 10 (2013): 198-199.

[21] This is also a philological consequence of the relationship between texts and the Web. From a philological perspective, this can be seen, for example, with regard to the apparatus of a critical edition. If in its paper form the latter can only indicate the variants in which the realized edition of the text differs from the tradition, the web, on the other hand, can make available all the witnesses of which the apparatus merely gives succinct references. It is therefore obvious that having made available the

thinks only of the annotations, endnotes, footnotes, dedications, prefaces, or appendices that would be inexorably erased with a digital transposition of only the textual core of a work. Instead, it is evident how materials of this kind are not ancillary to the text; on the contrary, they contribute to better qualify the integrity of its content. And yet, as noted by Elena Pierazzo, it is unmistakable that many modern digital formats have been designed in a different direction, the one that favors an exclusive, direct consultation of the central body of the digitized text, with respect to which the accompanying elements are sacrificed.[22] In order to forestall such a loss, researchers should seek to enhance even more the already excellent philological acumen and exegetical spirit that has guided the digitization tools and projects. Dante is undoubtedly an important test case in this regard, given the peculiar and fragile state of the manuscript tradition of his works. It is thus a logical next step that, in the near future, professional teams or crowdsourcing projects (perhaps in the latter case making use of Wiki platforms) could use digital tools to display and enhance the paratext of Dante's works. The hope is that future Dantean projects can reach the same quality that today distinguishes works such as the digital edition of the *Commedia* or the *Monarchia* edited by Prue Shaw, examples of true excellence in the Digital Humanities and role models even for those not involved in Dantean matters.[23] Such as these latest works,

---

integrity of the material to which the apparatus refers, one will tend to favor the set of texts poured online and not the mere informative list of them proposed by the apparatus. The latter will thus lose its centrality. Addressing the complex problem of the arbitrariness of the readings of the textual lessons of the *Commedia* Charles Singleton called for the necessary publication of a commentary on the poem that would disregard only one interpretation, but also make available those that have been discarded in preferring the printed one. Such a problem could be solved precisely by the digitization of the commentaries. The *Dante Lab Reader*, which allows simultaneous comparison of multiple commentaries on Dante's poem, partially obviates the shortcoming pointed out by Singleton.

[22] Elena Pierazzo, "Teoria del testo, teoria dell'edizione e tecnologia," *Ecdotica* 14 (2017): 135–148.

[23] See Paolo Chiesa, "L'edizione critica elettronica della Monarchia: la filologia informatica alla prova dei fatti," *Rivista di Studi Danteschi* 7 (2007): 325–354; Prue Shaw, *Commedia. A Digital Edition* (Florence: SISMEL, 2021); Gian Paolo Renello, "Appendix. On the programs used for the digital edition of the Monarchia and the Commedia," in *Everything you always wanted to know about Lachmann's method. A non-standard handbook of genealogical textual criticism in the age of post-structuralism, cladistics, and copy-text*, ed. P. Trovato (Padua: Libreriauniversitaria.it, 2014), 224–227; Vera Ribaudo, "Nuovi orizzonti dell'ecdotica? L'edizione elettronica della Monarchia e della Commedia di Prue Shaw," *L'Alighieri. Rassegna dantesca* 54 (2013): 95–127; Peter Robinson, "Electronic editions which we have made and which we want to make," in *Digital philology and Medieval texts*, ed. F. Stella, A.

~ 216 ~

digital editions that incorporate paratextual elements would have to reconcile critical expertise, pioneering technologies (e.g. the versioning software for comparing manuscript variants) and methodologies (such as photonics or image saturation control for the study of palimpsests) to safeguard every component of the text. Indeed, what Giorgio Petrocchi wrote in the celebrated essay *Radiografia del Landiano* (*Itinerari danteschi*, 1969) about the oldest preserved codex of the *Commedia* remains valid to this day: "[sono g]li strumenti scientifici d'oggi [che] consentono di ripresentare il problema del testo primitivo […] nella sua globalità, scalzando fin dove è possibile l'opera del raschino iconoclasta."[24]

While these may appear to be high-cost reconstructive interventions, which might redefine the logics, practices, and structures of digital Dantean applications, arrangements such as those I have suggested are as close as one can come to reconciling online realities with more traditional working environments. With measures such as these, the vision set forth by Francesca Tomasi might be realized: that "[a]rchivistica e filologia, digital humanities e semantic Web assieme a user experience e information visualization [possano cooperare] per fare delle risorse informative degli autentici knowledge sites."[25]

However, it is undeniable that there exists a certain resistance to implementing changes such as those envisaged. Such resistance primarily concerns limits of expertise and technical skills among humanists, in the absence of which becomes necessary a close and constant cooperation with professional computer scientists with well-defined task performance and possibly high level of daily

---

Ciula (Pisa: Pacini Ricerca, 2007), 1-12; Paolo Trovato, "La doppia Monarchia di Prue Shaw," *Ecdotica* 7 (2010): 193-207.

[24] Giorgio Petrocchi, "Radiografia del Landiano," in *Itinerari Danteschi*, ed. C. Ossola (Milan: Franco Angeli Editore, 1994), 134; Ezio Raimondi was far-sighted: "Certo, il filologo può continuare a occuparsi del proprio giardino, a misurarsi con i casi concreti, e sono quelli che contano, della sua disciplina nelle forme che le sono proprie da lungo tempo. Ma il solo fatto che esistano dispositivi tecnici quali i calcolatori elettronici, di cui egli può già disporre oggi per le concordanze di un testo, nell'ambito dell'applicazione più comune, costituisce un problema nuovo che non può essere risolto in un semplice giudizio d'uso, o di sfruttamento, del tutto esterno, e che in qualche modo, trascendendo la prassi, tocca l'idea stessa della filologia" [Ezio Raimondi, "La filologia moderna e le tecniche dell'età industriale," in *Tecniche della critica letteraria* (Turin: Einaudi, 1975), 91].

[25] Francesca Tomasi, "Edizioni o archivi digitali? Knowledge site e apporti disciplinari," in *Edizioni critiche digitali. Edizioni a confronto*, ed. C. Bonsi, P. Italia (Rome: Sapienza Università Editrice, 2016), 135.

~ 217 ~

interaction.[26] That is, we need, as Jeffrey Schnapp writes, a sharp Dante connoisseur and founder of Harvard's meta-LAB, "una literacy […] trasversale perché queste diverse expertise, radicate sempre in saperi settoriali, devono dialogare fra loro attraverso un linguaggio comune, capace di convertire e adattare saperi specialistici."[27] Demands of this kind can only be met if one has the support of entire research centers, digital specialists, teams of humanities researchers and, last but not least, both substantial funds to finance the activities and sufficient (and not insignificant) time investments on the parts of the PIs.

However, this does not detract from the fact that alternative solutions are possible, probably less sophisticated from a programming perspective, but equally functional for the refinement of research as well as consistent with the principle of IT affordability. It is precisely in keeping with such digital sustainability that the technological section of my doctoral project fits in, which has as its ultimate goal the coding and release of a new Dantean repository called the *Database Allegorico Dantesco* (DAD) and on which I will now focus my attention.

However, before proceeding to explain structure and functions of DAD it is good to conclude by recalling how everything that has been said so far should not be taken as a *vademecum* of universal applicability. That is, this is not to say that every Dantean site must necessarily conform to the suggestions presented thus far, nor that these resources are ineffective because they have not done so. In addition, as alluded to at the beginning of this paper, other corrective procedures are also certainly possible, the description of which would, however, find a place in a discussion devoted to the benefits that computer methodologies could bring to Dante Studies in general.

If a variety of methodical observations have been proposed, it has been done so because each Dantean website specializes in a

---

[26] Extremely pertinent is Paola Italia's observation regarding the shift in perspective that has come about with the technological hybridization of critical-philological work processes: "Al filologo digitale [ma il discorso può tranquillamente valere anche per un semplice critico della letteratura] […] con gli standard attuali, e in mancanza di un modello editoriale in cui siano automatizzati i processi di codifica, sono state richieste competenze tecniche che, quando non hanno scoraggiato i filologi, hanno sottratto tempo ed energie al lavoro di documentazione, di studio, di interpretazione. Che dovrebbe essere il cuore del lavoro filologico, spesso ridotto a una meccanica collazione di testi e codifica ("taggatura") di documenti" [Paola Italia, *Editing duemila. Per una filologia dei testi digitali* (Rome: Salerno Editrice, 2020), 180].

[27] Jeffrey Schnapp, *Digital Humanities*, ed. M.G. Mattei (Milan: Egea, 2015), 24.

~ 218 ~

particular area, addresses peculiar types of users, and pursues specific goals. The variety of advice is justified in light of the structural diversity of Dantean resources. It would thus turn out to be anachronistic to wish that each of these would adapt to the principles of the Semantic Web, those of close/distant reading, or be remodeled for paratextual searches. Instead, it was intended to show how there are possibilities and functions that while coherent with the evolution of information technology realities have not been taken into account – or not completely – by those concerned with Dante. In short, it was intended to prove how methodical alternatives to those now established in digital studies on Dante are possible and I also personally believe that by adopting them the benefits can be significant.

3. *Database Allegorico Dantesco (DAD)*

As an integral part of my doctoral research project as part of the course in Humanism and Technology funded by the University of Macerata (under the supervision of Prof. Laura Melosi) and co-directed with the University of Bologna (under the supervision of Prof. Giuseppe Ledda), the *Database Allegorico Dantesco* (DAD) intends to make available a vast selection of critical material pertaining to the theme of allegory in the *Divina Commedia* (fig. 4).

Fig. 4 – *Database Allegorico Dantesco*'s logo

The project has been conceived and programmed to serve both as an introductory study tool for neophytes of the issues dealt with in it, and as a tool for in-depth analysis of the poem's allegory for scholars already acquainted with its most complex features. Although we have chosen to limit the scope of the study to only three cantos of the *Commedia*, *Inf.* 9, *Inf.* 1, *Purg.* 8—selected for the strategic role they have assumed in relation to the discourse on allegory and allegoresis in the poem—we have also prepared a section in which will be collected and displayed the main works that

~ 219 ~

national and international *dantistica* has produced over the past 120 years on the topic of this research.

To date we have made available for free consultation more than 400 texts, totaling more than 7000 digitally transposed pages and referring to more than 250 different authors, freely available for consultation. A larger percentage of these represent Italian and English entries, with a smaller representation of French and Spanish exponents.[28] Here I should note that, in compliance with the copyright on published texts, all digital reproductions that exceed 15% of the analog material make reference to sources that are already freely and fully available online (falling under policies for fair use).[29] The database has several advantages. First of all, the disaggregation of data is obviated. In fact, there are no web aggregators that accommodate in the same place all the material present in the DAD. This precludes the major limitation of imposing on the user long, sometimes complex or convoluted—and, in some cases, no longer even feasible—searches across several web domains. In addition, the texts that are already available online do not allow the kind of interoperability I am proposing with the DAD.

Before briefly illustrating the basic architecture of the repository and discussing plans for its ongoing development,[30] it is important to remark how common computer logic was taken into account in the realization of the project. This choice was not trivial or superficial. On the contrary, it is essential to the site's basic functionality. Even before the long phase of configuring the database—selecting and retrieving the texts, manually transcribing them into a machine-operable format, and then populating the database—
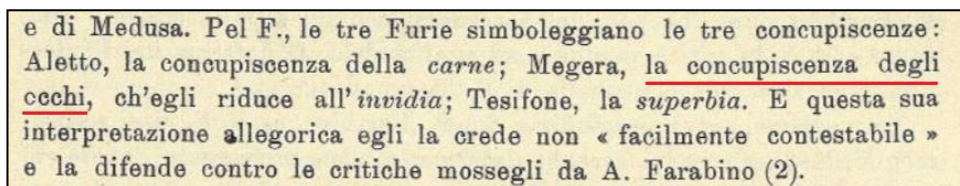
---

[28] The data shown here are as of October 2022.

[29] "La Direttiva europea per il copyright consente la riproduzione per finalità di conservazione di tutte le opere, senza considerare se essere siano coperte da diritti e senza il requisito preliminare del consenso dell'autore. L'art. 5.2.C di tale Direttiva prevede un'eccezione a favore di archivi o biblioteche pubbliche, istituzioni accademiche o musei, che possono svolgere specifiche campagne di riproduzione senza scopo commerciale" [Maurizio Borghi, Stavroula Karapapa, "Dal cartaceo al "digitale di massa." Biblioteche virtuali, diritto d'autore e il caso Google Books," in *Teorie e forme del testo digitale*, ed. M. Zaccarello (Rome: Carocci Editore, 2019), 106]; see also Diana Kichuk, "Loose, Falling Characters and Sentences. The Persistence of the OCR Problem in Digital Repository E-Book," in *Libraries and the Academy* 15 (2015): 55-91.

[30] For a more in-depth discussion of the same, see Matteo Maselli, "Struttura e funzioni di un database dantesco. Proposta per un'inedita risorsa digitale," in *L'altro Dante. Processi di attualizzazione della Commedia. Atti del Seminario Internazionale di Studi 24-25 novembre 2021* (Lanciano: Carabba, 2022), 241-260; Matteo Maselli, "Database Allegorico Dantesco," *Umanistica Digitale* 13 (2022): 165-170.

~ 220 ~

there was an equally intense period of evaluation at the end of which we decided to favor, as the basic principles of the entire work, perspectives that were in line with the everyday use of the most common technological resources. Analysis of this kind, coherent with the models of digital sustainability mentioned earlier, were part of the "modeling" phase of the DAD. That is, during this preliminary moment, each object to be included was screened, its main characteristics and functions were chosen, and aspects that were judged to be impractical or not particularly centered with the objectives of the research were discarded. As a rule, we sought to adhere to the so-called "model-for" theorized by Willard McCarty, an operational protocol by which one aims to produce something new and not simply to describe something already existing (what McCarty calls a "model-of").[31] During the initial evaluation of the tool, it was possible to see the validity of McCarty's iterative perspective: we realized that there is a need for a progressive trial stage and constant employment of the product in order to grasp its limitations and the areas that most urgently require corrective action or additions.[32]

This preparatory phase informed the logical principles chosen for the DAD construction, which, as will be shown, concern hypertextuality between archived texts and correlation between data. To make these operating criteria clear, each text has been fully rewritten and marked up in XML (eXtensible Markup Language), carefully following the guidelines set by the Text Encoding Initiative (TEI).

e di Medusa. Pel F., le tre Furie simboleggiano le tre concupiscenze: Aletto, la concupiscenza della *carne*; Megera, la concupiscenza degli occhi, ch'egli riduce all'*invidia*; Tesifone, la *superbia*. E questa sua interpretazione allegorica egli la crede non « facilmente contestabile » e la difende contro le critiche mossegli da A. Farabino (2).

Fig. 5 – Among the limitations that threaten the accuracy of OCR software operations the most common is the altered perception of the letters of a term, a condition that results in a departure from the objectivity of the message proposed by a text. In the case reported here, the inaccurate graphical rendering

---

[31] Willard McCarty, *Humanities Computing* (Basingstoke-New York: Palgrave Macmillan, 2005).

[32] At first, for example, the possibility of inclusion, as will be discussed shortly, of underlining and annotation functions by the user was not taken into account. Later, however, after an initial phase of use, it proved almost natural to enclose them so as to increasingly thin the differences between a paper text and its digital counterpart.

~ 221 ~

of the characters "o" and "e" of the word "occhi" led to its scanning of the lemma as "cechi," thus reversing the meaning of the entire passage

To complete the XML transcription, we first had to create digital copies of the texts. We did so using Optical Character Recognition (OCR) scanners, which created text files that could then be subjected to manual correction through word-processing programs. Even automatic digitization is in fact not free from error.[33] Inaccurate reading of textual gaps, graphic misunderstandings or misidentification of print characters are statistically very frequent limitations in automatic transposition processes (fig. 5), and therefore the only way to achieve a highly satisfactory level of accuracy is through manual correction at the post-production stage of digitized texts.

Having cleaned the raw OCR texts and properly carried out the rewriting in XML, a final format was obtained that presents advantages not only in terms of preservation, but also and above all from the point of view of the possibility of use.[34] The nested structure of XML text allows for the insertion of additional particles (tags) that attribute further characteristics to the textual area in question.[35] Through tags the text is rendered more dynamic and responsive to external requests, such as the forwarding of targeted searches. In the extensive taxonomy of tags, we deemed it especially advantageous to make use of the marker <ref target="…"> … </ref>. Its inclusion in the schema of the XML transcription made possible interactive intertextual relationships in the catalogued documentation, thus recreating that condition of textual consultation recommended for any semantically complex author, and which in the case of Dante has even been codified as an official

---

[33] For example, for a printed text we need to "procedere al riconoscimento e alla divisione della pagina di sinistra e della pagina di destra (splitting), alla correzione dell'eventuale rotazione (deskewing) della pagina, al riconoscimento degli specchi di testo sulla pagina (content box recognition), all'eliminazione di macchie (despeckling), all'eventuale stiramento dello specchio di testo (dewarping) e infine alla binarizzazione (binarization) con i valori ottimali di luminosità e contrasto" [Federico Boschetti, *Copisti Digitali e Filologi Computazionali* (Rome: CNR Edizioni, 2018), 13].

[34] Since a standard format was used for the DAD texts, it will be possible to make use of them on multiple devices (PCs, smartphones, tablets) and possibly transport archived data to other web infrastructures. Such arrangements are part of the always necessary attempts to counteract computer obsolescence.

[35] See Susan Hockey, "The Reality of Electronic Editions," in *Voice Text Hypertext*, ed. R. Modiano, L. Searle (Seattle: University of Washington Press, 2003), 361-377.

exegetical methodology.[36] Thus, the user consulting the DAD files[37] will find hyperlinks, corresponding to direct quotations in the body of the text or indirect bibliographical references in footnotes, that direct users to the works mentioned, and which can thus be immediately examined with a click (fig. 6).

```xml
<note place="bottom">Tra gli studi del Singleton non divulgati in versione italiana, ricordiamo
  <ref target="https://www.databaseallegoricodantesco.it/view.php?id=446">The irreducible Dove,
    in «Comparative Literature», IX (1957), 2, pp. 132-133</ref>, che è risposta a
  <ref target="https://www.databaseallegoricodantesco.it/view.php?id=344"> R.H. GREEN,
    Dante's «Allegory of poets» and the Medieval Theory of Poetic fiction nella stessa rivista,
    pp. 118-128)</ref></note>.
```

Fig. 6 – Example of a footnote written in XML and containing hyperlinks to works stored in the DAD

---

[36] I refer to the pervasive practice of "leggere Dante con Dante": "[a] me pare […] che per giungere all'intelligenza della polisemia dantesca, occorra seguire il Dante 'secondo Dante,' che è l'insostituibile *autoaccessus* alla propria opera, altrimenti impossibile da cogliere. Da qui, a sostegno e ad illuminazione, l'incessante cura autoesegetica" [Vittorio Cozzoli, "Il fondamento della polisemia dantesca," in *Lectura Dantis 2002-2009. Omaggio a Vincenzo Placella per i suoi settanta anni* (Naples: Università degli Studi di Napoli "L'Orientale" 2011), 827]; On the same subject, see at least Gianbattista Giuliani, "Dante spiegato con Dante," in *Dante e il suo secolo* (Florence, 1865); Edward Moore, *Studi su Dante*, ed. B. Basile (Rome: Salerno Editrice, 2015); Emilio Pasquini, *Intertestualità e intratestualità nella Commedia dantesca. La tradizione del Novecento letterario* (Bologna: CUSL, 1993). Benedetto Croce did not shy away from the issue either, albeit only alluding for the *Commedia* to a practice that Dante himself, had he had the opportunity, would likely have, as in part he actually did, practiced: "Cominciò questa interpretazione filosofica ed etica e religiosa [spiegare Dante con Dante] fin dai tempi di Dante […]; e sarebbe probabilmente cominciata per opera sua stessa, se gli fosse bastata la vita, perché chi aveva commentato le proprie canzoni nel *Convivio*, difficilmente avrebbe lasciato senza chiosa il 'poema sacro'" [Benedetto Croce, "Introduzione," in *La poesia di Dante*, ed. G. Inglese (Naples: Bibliopolis, 2021), 10-11]; It is also important to keep in mind Jean Pépin's observation: "[…] in Dante scopriamo, accanto a un'esegesi allegorica che prende come oggetto i dati a lui esterni, un'altra esegesi del tutto simile che egli saggia sulle sue stesse opere. Se per designare quest'ultimo procedimento osassimo forgiare un neologismo […] potremmo parlare di 'auto-allegoresi'. Il termine […] corrisponde a un aspetto originale della pratica letteraria di Dante, che appare senz'altro estranea ai suoi predecessori e ispiratori, tanto patristici che medievali" [Jean Pépin, "Lettera e allegoria in Dante," in *Leggere Dante. Antologia della critica dantesca*, ed. A. Russo, E. Schiavina (Bologna: Zanichelli, 1976), 163].

[37] For the online upload of the DAD cards the source files in XML (input) were transposed, through a specific editor (oXygen was used in this case), into a digital output (output) object in HTML.

~ 223 ~

Simultaneous comparison makes available multiple critical voices on the object of research, thus allowing for an overall evaluation that also takes into account positions contrary to the dominant ones. In this way, the DAD encourages the emergence of an extensive exegetical network related to Dante's allegory.[38] For such a function to be guaranteed it is necessary to make sure that the text called into question by a given author has been duly acquired and placed online in the database. This involved a careful choice of parameters for retrieving archived material.

In the complicated selection process—a difficulty due as much to the extensive chronological span taken into account as to the even larger bibliography available—priority was given to the authority of a particular study and/or its author and to the high rate of recursiveness of texts between independent works due to direct and/or indirect citations. Critically important and historically recursive texts are indeed particularly prone to be subject to intertextuality.

Hypertextuality also affects the indexes where digitized texts are composed of multiple internal units (chapters, paragraphs, etc.). The interactive table of contents, with its direct links, makes it possible to arrive automatically at the various subdivisions.
Regardless of the type of work chosen, each database card will then display the metadata, the text, and related notes. Such informational prospectus is necessary because, as Peter Shillingsburg has argued, any digital reproduction is always derivative, and thus it is inadmissible to omit or ignore the mode of reproduction of a text and the basic information for their unambiguous identification.[39] Hence we deemed it necessary to individuate the categories of metadata, text, and notes.

---

[38] Outside of a purely computerized dynamic, the idea that allegory was part of and in turn contributed to determining a complex system of thought, extensive and constantly expandable to the point of encompassing within itself, much like a digital network, a plurality of issues was a firmly held belief in medieval times. Indeed, allegory was not held to be a mere rhetorical tool but raised to a means for the attainment of metaphysical knowledge according to rules in which progressive reference to other concepts was a common perspective. Typical is the Eriugena image but made known in Dante studies especially by the Boccaccian *Tattatello in laude di Dante*, of the peacock's tail which, though unitary, contained within itself, in each of its individual points, a cross-reference to an uncontrolled variety of colors – precisely a visual-conceptual network [John Scotus Eriugena, *Periphyseon*, III, P.L. 122, 749C; see also Giulio D'Onofrio, "Allegoria e metodo," in Id., *Per questa selva oscura. La teologia poetica di Dante* (Rome: Città Nuova Editrice, 2021), 57].

[39] Peter Shillingsburg, "From Physical to Digital Textuality. Loss and Gain in Literary Projects," *CEA critic* 76 (2014): 158-168.

~ 224 ~

Because one of the main criticisms levelled at digital archives is that they often downplay a complete and clear bibliographical overview of the texts preserved, in the metadata area the DAD offers due bibliographical information, allowing for the unambiguous recognition of the consulted work.[40] Then we have the body of the text itself, which can in a sense be defined according to a diplomatic perspective, as it corresponds exactly to the original physical product. Finally, footnotes will be located at the end of the page and will be accessible with interactive pointers (fig. 7).

Fig. 7 – Example of DAD card with highlighting of table of contents, metadata, text, and notes

In order to make the experience of consulting the database cards close to that which should normally distinguish a knowledge site, we decided to include specific options in order to enhance the

---

[40] See Harald Brandes, "Problemi tecnici attuali e futuri degli archivi esistenti e di quelli in via di realizzazione," in *L'eclisse delle memorie*, ed. T. Gregory, M. Morelli (Rome-Bari: Editori Laterza, 1994), 33-43; Kenneth M. Price, "Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?," *Digital Humanities Quarterly* 3 (2009); Kate Theimer, "A Distinction Worth Exploring: Archives and Digital Historical Representations," *Journal of digital humanities* 3 (2014).

~ 225 ~

ways in which the collected works could be read. Aware of the possible benefits that the study of the texts could receive from the freedom of highlighting and commentary, similar with those that distinguish an analog book, underlining and annotation functions were included, easily performed by resorting to a special toolbar located at the top of each card (fig. 8).

Fig. 8 – Toolbar within the DAD cards through which to select various text marking options (highlighting and/or annotation)

To avoid the invasiveness of such arrangements—which could have changed the basic format of the texts on which the user will have autonomy of intervention—each user will be able to create a personal profile in which to work directly on the texts and automatically save the changes made without altering the appearance of the originals.

Slightly more complex are the arrangements required to follow up on the principle of correlation, another protocol followed for DAD programming. By this, I mean the software's automatic suggestion of items not considered by the user in the search phase. The software will in fact draw alternative keywords from a list of related terms compiled on the back-end page.

Suppose that in a *lectura Dantis* from *Inf.* 9, digitized and entered into the DAD, a Latin-educated commentator discussed Megaera, Alecto and Tisiphone, the three monsters who invoke Medusa to stop Dante's way, resorting only to the terminology of Roman mythology that identified the three characters as "Furie." It is obvious that because of this lexical choice, should one go to search for the term "Erinni," synonymous with the Furies but of Greek descent, the software would not return the *lectura Dantis* in question. Since such a selection burdens the quality of the search conducted by the user, who in all likelihood also has an interest in analyzing texts in which "Erinni" is mentioned, expedients have been devised to preclude the omission of meaningful data. On the back-end of the database (which, to protect the integrity of the resource, can be accessed only by the site administrator by logging in with personal credentials), a table of related searches has been set up in which relationships between terms have been established, by which we aim to automatically solve the problem of omission in

~ 226 ~

the results indexed by the software (fig. 9). This is a publishing infrastructure developed by resorting to a specific framework but dependent on a logic common to every web user's experience (we can think for example to the suggestions offered by Google when the name of a search term is typed incorrectly).

Fig. 9 – Table of related searches present in the back-end area of DAD

When the admin user completes the form with appropriate search parameters and related terms—an operation that can be expanded without limitations, even to several simultaneous terms—the software will propose suggestions of alternative entries to the one searched. By clicking on the suggested terms, the user will be automatically redirected to the complete list of DAD texts that contain them.

The search engine has also been set to allow it precise recognition of the value of one or more keywords entered in the query mask, so as to ensure absolute relevance of the outcomes returned by the machine. Thus, if, for example, the term "metafisica" is entered, one will be referred only to the cards that contain its exact wording, automatically omitting those in which only "fisica" is mentioned, those same cards which would have been proposed if the input-output relationship of the software had instead been calibrated on the identification of only the letters of the headwords ("fisica" being contained within the term "metaFISICA").

Lastly, one can select the path in which to conduct the investigation and decide whether to expand it to the entire DAD corpus or to limit it to specific categories/subcategories, with the premise however that all texts, with their places of preservation, loaded into the database will be appropriately indicated in a separate section ("Archive"). In fact, for a thematic order that would facilitate the user's orientation, four sections for the collection of the

~ 227 ~

digitized material have been specially arranged, internally composed of further partitions: Historical Exegesis; *Lecturae Dantis*; Essays and scientific articles; Monographs. The first three concern the cantos selected for the project, while the fourth pertains to the best outcomes of research on allegory in the *Commedia* published in the 20th and 21st centuries. Pre-twentieth-century studies, such as the interpretations of the cantos made from the earliest Dantean commentators, will be located in a special section on historical exegesis.

I would conclude by recalling again the project's nature *in fieri*, and how, therefore, changes in content and structure, even major ones, cannot be ruled out in the near future. I hope I have made clear, however, that if such changes are applied, it will be exclusively to foster a better characterization of the resource, always taking into consideration the needs of the users it addresses.

4. *Conclusion*

With the present paper I have sought to propose, first theoretically with a review of possible logical expedients, and then practically with the presentation of the *Database Allegorico Dantesco*, development methodologies that can enhance the opportunities for study currently offered by the considerable series of Dantean sites. The latter do not lose their attractiveness and efficiency since, if well employed, can facilitate the management and analysis of Alighieri's texts. Moreover, their proper use not only enables one to obtain in-depth knowledge about Dante, but even to redefine the researchers' *modus operandi*. Therefore, it is evident how even the smallest and seemingly insignificant design flaw can compromise the functionality of these resources and the quality of the research done using them. This may limit the hermeneutic or reading perspectives of those who rely on such tools to the point of preferring them to a traditional paper medium.[41] Such a choice, therefore, obliges a level of constant monitoring that can intercept and correct systemic flaws that would negatively scale back the operation of the machine to which we increasingly reserve "quella fede che vince ogne errore" (*Inf.* 4. 48).

---

[41] Carol Chiodo, "Beatrice in the Tag Cloud," in *Approaches to Teaching Dante's Divine Comedy*, ed. C. Kleinhenz, K. Olson (New York: Modern Language Association of America, 2020), 257–261; Laura D'Angelo, "Dante tra web e social network," *Studi Medievali e Moderni* 25 (2021): 699–717; Crystal Hall, "Digital Humanities and Italian Studies. Intersections and Oppositions," *Italian Culture* 37 (2019): 97–115.

~ 228 ~