

The Digital Humanities Classroom as a “Node”. From Toolbox to Mindset?

Florentina Armaselu

Luxembourg Centre for Contemporary and Digital History
(C²DH), University of Luxembourg

The paper describes a framework for digital pedagogy and evaluation that combines different approaches – theoretical, application-oriented and project-based –, underpinned by a view of the digital humanities classroom metaphorically defined as a “node”. The metaphor encompasses the hybrid character of digital humanities teaching, as a crossing or convergence of paths, which involves the use of various categories of study materials, themes, examples, methodologies and skills, as well as exchange, circulation of ideas and connection with other disciplines. It is assumed that the application of this type of hybrid approach in the course design, together with the coverage of a variety of tools to be studied in class, might foster a reflective mindset, enabling the students to contextualise the digital technologies within different humanities areas, beyond the acquisition of technical skills and knowledge of computational methods and tools usually referred to as the toolbox-based education paradigm.

1 Introduction

This paper is based on the case study of an introductory course on computational text analysis and text interpretation for students enrolled on the Bachelor en Cultures Européennes (BCE) programme at the University of Luxembourg. Given that digital humanities (DH) education is a hybrid field, involving various subject matters and skills as well as teaching and assessment strategies, the paper will use the lens of a classroom case study to address the challenge of DH teaching, with a shift from a toolbox-based instruction paradigm to a more reflective mindset that considers new technologies in a broader cultural and pedagogical context.

Studies in digital pedagogy have already pointed out the need for “curriculum change” and a shift towards an “inquiry-based curriculum” Whitby (2007, pp. 3, 4) fostered by new technologies and focusing less on “things to know” and more on developing “strategies for learning” by stimulating the imagination and engaging students in “taking responsibility for their own learning” (pp. 7, 8). By questioning the “what works” paradigm “currently dominating educational research”, Ross (2017, pp. 1, 2) assumes the relevance of “speculative methods” that may capture the “‘not-yetness’ of technologies” and provide a “conceptual handle for digital education

approaches” and for openness to new practices and ideas while allowing “curiosity, critique, doubt, unintended consequences and emergent properties of technologies in use”. Other studies, focusing on digital humanities, have highlighted different aspects to be considered when teaching in this area. Mahony and Pierazzo (2012, p. 7) argue that DH teaching should be relevant to the student’s study and research interests and should deal not so much with skills, though these are important, but with “new methodologies and new ways of thinking”. Cordell (2019) recommends cultivating a “mindset for approaching data, exploring it” and understanding “what questions computation might help answer about it”, rather than aiming to bring the students to expertise in any particular computational method. In their article “Beyond buttonology”, Russell and Hensley (2017) affirm that tutorial-based teaching focusing on tools should be complemented by “critical engagement with digital methodologies” and “humanities sources as data”. Other scholars (Papadopoulos and Schreiber, 2019; Sinclair and Rockwell, 2012) advise practical approaches in the classroom, such as asking students to develop their own research questions based on their own texts or applying problem-/project-based learning and enabling students to become producers rather than just consumers of knowledge.

The main questions derived from these readings that the paper will address are the following. (1) What type of information and pedagogical approach should learners be exposed to in order to attain the objectives of such a shift? (2) How effective is the applied methodology in attaining these goals? The remaining sections will present in more detail the initial settings, assumptions and methodology of the course, as well as the results and feedback obtained from the first (and partially second) iteration of this empirical case study.

2 Case study

The course was taught during the 2019-2020 winter semester, starting in September 2019 and ending in February 2020 with a final examination. It was an optional general course worth three ECTS¹ in the Arts and Media Studies module of the BCE programme. Eleven undergraduate students completed the course, with profiles in different areas such as history and English literary and linguistic studies. The course, Introduction to Computational Text Analysis and Text Interpretation, introduced concepts, methods, tools and data for computational text analysis and interpretation from six categories: corpus linguistics, parsing, named entity recognition, sentiment analysis, topic modelling and word embedding. Based on an application-oriented approach, it was intended to illustrate how these techniques can be applied when answering or formulating historical, linguistic or literary research questions, for instance when analysing the evolution of a concept in an art history collection, main topics in the European Parliament’s news releases, specific vocabularies in a series of transcribed life-history interviews or the sentiment-based plot arc of a novel. At the same time, the course aimed to foster a basic understanding of the theoretical assumptions underpinning the “black box” guise of the user interface and assimilation of elementary principles of programming in R and Python. Each session, except for those dedicated to work on the project, included theory and a presentation of potential applications followed by hands-on activities. The final assessment consisted of individual projects, with the students using the tools of their choice from those studied in class.

¹ European Credit Transfer and Accumulation System. See https://ec.europa.eu/education/resources-and-tools/european-credit-transfer-and-accumulation-system-ects_en.

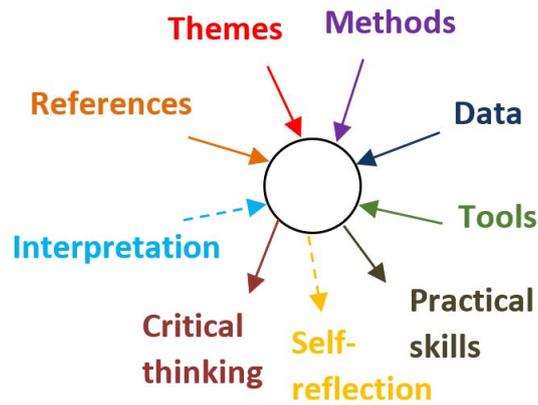


Figure 1: The DH classroom “node” metaphor.

Taking the node or crossroads metaphor as a starting point (symbolically represented in Figure 1), the pedagogical approach applied in the classroom supposed the intersection of different areas of enquiry and virtual connections within the course or programme as a whole, including for instance thematic (culture-oriented *themes*), theoretical (text analysis *methods*), data-related (*dataset* processing and assessment), practical (hands-on activities with different *tools*) and referential (*further reading* and *links* to relevant information to be studied individually) aspects.

Expected outcomes included *critical thinking* and *practical skills*, enabling the students to apply the knowledge acquired to their course project or projects from other disciplines. The dotted aspects resulted from the study and will be discussed in section 3. These assumptions were tested via an anonymised questionnaire proposed to the students and completed at the end of the course, which included questions such as: (1) the role of the selected text analysis tools in their projects and in answering the research questions; (2) the “added value” determined by the use of this type of analysis as compared with a more “traditional” non-digitally-based study and interpretation of texts; (3) reflections on the innovative character (if any) of the computational text analysis approach and/or its limitations, bias, etc.; (4) evaluation of the course scenario.

3 Discussion

In his model of the public image of a city, Lynch (1997, p. 47) defines the concept of nodes as “primarily junctions, places of a break in transportation, a crossing or convergence of paths, moments of shift from one structure to another. Or [...] simply concentrations, which gain their importance from being the condensation of some use or physical character, as a street-corner hangout or an enclosed square.” This definition conveys a composite image, of foci to and from which one can travel, junctions, moments of shift, crossing of paths and concentrations, which served within the methodological framework of the course as a metaphor for the digital humanities classroom as a “node”. The metaphor encompasses the hybrid character of DH teaching, which involves a concentration of various subject matters, teaching and assessment strategies to be applied in the classroom, as well as circulation of ideas, exchange and connection with other disciplines inherent to the concept of node as a unit within a network. It was assumed that applying this type of hybridity as an underlying principle in the course design might foster reflection on the role and characteristic

usage of digital technology in the humanities, beyond the acquisition of technical skills and knowledge of computational methods and tools.

3.1 Lesson structure

Following this principle, the lessons were conceived as a combination of basic theoretical information about the tools and methods to be studied, together with topics and sets of data from different areas of enquiry (history, literature, philosophy, arts, etc.). These themes and datasets were used as examples and starting points for experimenting during the hands-on activities. The main goal was to provide an overview of different categories of tools and methods for computational text analysis and a thematic basis for interpretation via these tools. Six categories of digital tools (see section 2) were presented and worked with during the course sessions, which included lexical, syntactic and semantic perspectives considered as basic standpoints in analysing texts using computational methods. The themes were chosen to cover areas of interest for students with different backgrounds enrolled on the BCE programme and to provide a relevant context for applying what was learned in class. The predefined themes were also intended to offer some guidance and models to the students for formulating research questions and interpreting the analysis results in their own projects.

The lessons involved various degrees of heterogeneity. Some dealt with the investigation of specific concepts, such as culturomics (Michel et al., 2011), hermeneutical tools (Rockwell and Sinclair, 2016), distant reading (Moretti, 2013) and Vonnegut's (2010) chalkboard shapes of stories. The application of these concepts was studied using tools for n-gram detection and visualisation (Google Books Ngram Viewer), Web-based text reading and analysis (Voyant), topic modelling (MALLET) and lexicon-based sentiment analysis (sentimentr, syuzhet via RStudio ²). Other lessons relied on theoretical subjects from various areas, such as Pennebaker's (2013) linguistic features for "predicting" honesty and deception, P. Thompson's (2004) model of life-history interviews with black immigrants to Britain, Kripke's (1981) philosophy of naming, J.B. Thompson's (2007) study of ideology in modern culture or Hazard's (1961) reflections on the origins of contemporary Europe. These theoretical aspects were combined with methods and tools such as keyness-based comparison of two corpora (AntConc), textometry (TXM), named entity annotation and query (GATE), context-free grammar and dependency parsing (Stanford parser) and computation of word similarity and representation of meaning through vector semantics (gensim – word2vec via Python IDLE ³).

The datasets proposed for exploration or analysis were based on online sources, such as CLARIN Resource Families, Project Gutenberg, the Digital Corpus of the European Parliament (DCEP), Digital Humanities Resources for Project Building — Data Collections & Datasets and researchers' published data, and included a variety of textual collections, from oral history interview transcriptions, parliamentary debates, film summaries and novels to fake and real news, children's literature and academic articles in art history. The aim was to provide the students with a variety of materials (theory, tools, themes and data references), which could help them in formulating their own research questions and interpretation hypotheses in the development of their projects. More details of the lesson structure and short overviews are presented in the following examples. (1) *Ideology and grammar. Parsing* combined theoretical

² <https://www.rstudio.com/>.

³ <https://www.python.org/>.

enquiry referring to J.B. Thompson's (2007) "modes of operation of ideology" and Bourdieu's (2005) insight into language and "relations of symbolic power" with the use of parsing tools (Stanford parser) to analyse passivisation and nominalisation patterns in a sample of children's literature. (2) "*Distant reading*" of parliamentary discourse. *Topic modelling* elaborated on Moretti's (2013) and Underwood's (2019) reflections on close and distant reading, to experiment with topic modelling tools (MALLET) and discover hidden thematic structure and exploration paths in a sample of parliamentary news. (3) *Crisis and culture. Word embedding* started with Hazard's (1961) concept of "crisis of the European conscience" in the late 17th and early 18th centuries, opposing classical and modern views, and made use of word embedding techniques (gensim – word2vec) to detect context-based similarity for "crisis" in a corpus of academic papers on surrealism.

3.2 Assignments

The course included 14 sessions of 90 minutes each, 11 for subject presentations and hands-on activities and 3 for work on individual projects in a lab-like mode. Consultation hours and extra time to access the course computer room (4 additional sessions in the exam preparation period) were also provided for experimentation and finalisation of project reports. The final grade was based on five assignments: (1) project proposal, a 1-page term paper outlining the intended topics, research questions and envisaged approach – tools, methods and data to be used (30%); (2) project realisation, which consisted of the actual work on the project using methods, tools and data learned in class, and a set of deliverables such as input/output data samples, pieces of code, readme files, diagrams, etc. (30%); (3) project report, a 3-5 page written essay (20%); (4) a 10-minute oral presentation during the final examination that took the form of a mini-conference (10%); (5) active participation in the course as a whole (10%). The sixth assignment, non-graded, consisted of the students' involvement in providing responses to a feedback questionnaire. The first assignment (project proposal), which involved selecting methods and tools to be used in the project, was due after the 11 sessions of subject matter presentation (second week of December). Adjustments to these choices were accepted without penalties for slight changes between the proposal and the project report, since it was considered that methodology adjustments were part of the research process and work on the project. The other assignments were submitted at the end of the exam preparation period (end of January), with the final examination and oral presentation taking place at the beginning of February. Different types of documents to be submitted for the assignments were chosen (overview of the general idea, raw data and result files, essay, presentation, questionnaire answers) in order to document the process at different stages of the project and to enable reflection and a diversity of forms of expression as learning experience.

The proposed projects varied in terms of topics and selected tools. The majority of students (6) opted for the use of Voyant, some of them (3) in combination with AntConc for methodological comparison. GATE was chosen by a smaller number (3), while MALLET (1) and RStudio, sentimentr and syuzhet (1) were each selected by a single student. According to some of the reports, Voyant was preferred as it was considered to be accessible online, "easy to use", "user-friendly" and a "powerful tool for beginners and for advanced users". One student described Voyant as a "perfect tool to conduct a preliminary analysis of any text" but concluded that the "interpretations themselves would have to be tested for correctness using a different tool". No project

used the other remaining tools studied in class (Google Books Ngram Viewer, TXM, Stanford parser and word2vec via Python IDLE) in the first iteration,⁴ although some of them were mentioned as possible alternatives in the initial project proposals. From the point of view of themes and datasets, the projects dealt with various subjects at different scales, from the analysis of a small number of documents, e.g. specific US and European Commission presidential speeches, national and international treaties and conventions, and a novel, to the compilation and study of larger corpora containing speech transcripts of candidates in the 2020 US presidential race, extracts from the UK Hansard parliamentary reports, customer reviews from the Facebook page of a commercial clothing company or comments scraped from a dedicated pet discussion board. Regarding the research questions and the proposed approach, the students were interested in aspects such as: (1) comparing or identifying the development over time of different types of discourse (speeches, treaties, parliamentary debates) using corpus linguistics methods; (2) applying named entity recognition techniques to assess the capacity of the tool to detect unfamiliar names of people and places in a novel; (3) using topic modelling to identify the main arguments in the discourse of the top five Democrat frontrunners in the US presidential election; (4) computing sentiment scores for customer reviews and evaluating the overall accuracy of the method; and (5) creating a classification of pet-related subjects (categories, habits, needs) as reflected in the discussion forum of a pet-owners' community.

Although the variety of project ideas, themes, analysed data and overall results was generally rewarding, certain aspects, as discussed below and in the following section, required further attention and adjustments in a second iteration of the teaching experiment. Some projects showed initiative in collecting and preparing the sets of data, creativity in combining different features, tool tuning, labelling, visualising and interpreting the results, as well as awareness of the benefits and limitations related to the data size and format, tool, methodology and applied approach. Other projects, though demonstrating a relevant degree of effort in testing, producing raw results and interpreting, were less effective in communicating, through the final report and presentation, how the tools and methods were actually used to obtain the results and support the proposed interpretation. Additional categories included projects that, despite interesting thematic choices, focused more on the description of technical details of the analysis and less on the text interpretation itself, or projects that proposed interpretations mainly based on features that didn't go beyond lists of frequency counts, word clouds and word trends within documents. The latter cases were characterised either by over-interpretation of certain word occurrences within a given historical context or by a lack of contextualisation and simple linguistic comparisons of word usage. Although the general outcome was rather positive, there was an indication that the interpretative aspect of the node-based methodology (Figure 1, dotted input) was not sufficiently developed during the subject presentation and hands-on sessions of the course in such a way as to encourage adequate reflection and critical engagement with the results of the analysis. Correlating these observations with the students' feedback helped in evaluating the overall experience of the course and identifying some areas for possible improvement.

⁴ These tools, except for Google Books Ngram Viewer, were chosen in the second iteration of the course after some modifications of the lessons, as described in section 3.3.

3.3 Feedback

Feedback was provided via two channels: the official evaluation carried out at the Faculty level, after the end of the teaching period, and the responses to the questionnaire⁵ proposed as a sixth assignment at the end of the course, after the final examination. While ten (out of eleven) students completed the official evaluation, only eight also filled in the assignment questionnaire. No profile information was available in the official report. The proposed course questionnaire comprised a section for the description of the respondent profile, an anonymisation code and a formal agreement for the use of the collected data for research and publication purposes. The group that filled in the course questionnaire included 2 female and 6 male students, aged 18 to 34, enrolled on the BCE programme with a main background in history (4) and English studies (4). No previous knowledge of computational text analysis tools was reported. Six and two students selected answers 1 and 2 respectively on the Likert scale of 1 to 5 (*Not at all* to *Expert*) for the self-evaluation of their general knowledge of digital tools and methods. One student indicated Microsoft Excel as a tool already worked with before its use in class for the course.

Responses in the positive range were provided regarding the significance of the role played by text analysis tools in addressing the research questions formulated in the projects. Two answers were placed in the middle (3 points) and six answers on the right side (four/two respectively for 4 and 5 points) on the scale of 1 to 5 (*Not at all significant* to *Essential*). All the respondents agreed that these tools allowed them to discover something new or formulate new questions for the studied dataset in addition to their initial assumptions from the project proposal. Six out of eight also considered there to be an “added value” conferred by the use of this type of analysis as compared with a more “traditional” non-digitally-based study and interpretation of texts. Asked to provide details about this “added value”, the respondents mentioned the possibility of processing “large quantities of text in a short time”, providing a “quick overview [of the] main topics of a big text”, enabling “objectivity and orientation for textual analysis” or allowing the student to use these tools for “other classes and make [his/her] research easier”.

The questionnaire also included a section for the evaluation of the course scenario (Figure 2). The overall assessment ranged from four answers (one/three) in the *not interesting* and neutral area to four (three/one) in the *interesting* and *very interesting* area. The theoretical materials were considered *appropriate* and *very appropriate* by six students (three/three), while two (one/one) rated them as neutral or *not appropriate*. Five students (three/two) *agreed* or *fully agreed* that the hands-on activity provided enough background for the development of the projects, while three (two/one) answered either neutrally or *didn't agree*. Regarding the course assignments, one student evaluated them as *very difficult*, three as neutral and four (two/two) as *moderately* or *not at all difficult*. Although the number of respondents was relatively small, the answers provided a rough idea of how the pedagogical approach was perceived by the students.

⁵ The questionnaire and students' responses can be consulted as supplementary material (Excel file BCE-EU-105_02_CourseEvaluationQuestionnaire) published with the first version of the paper on Zenodo: <https://zenodo.org/record/3872252#.YKS8qKgZ2x>.

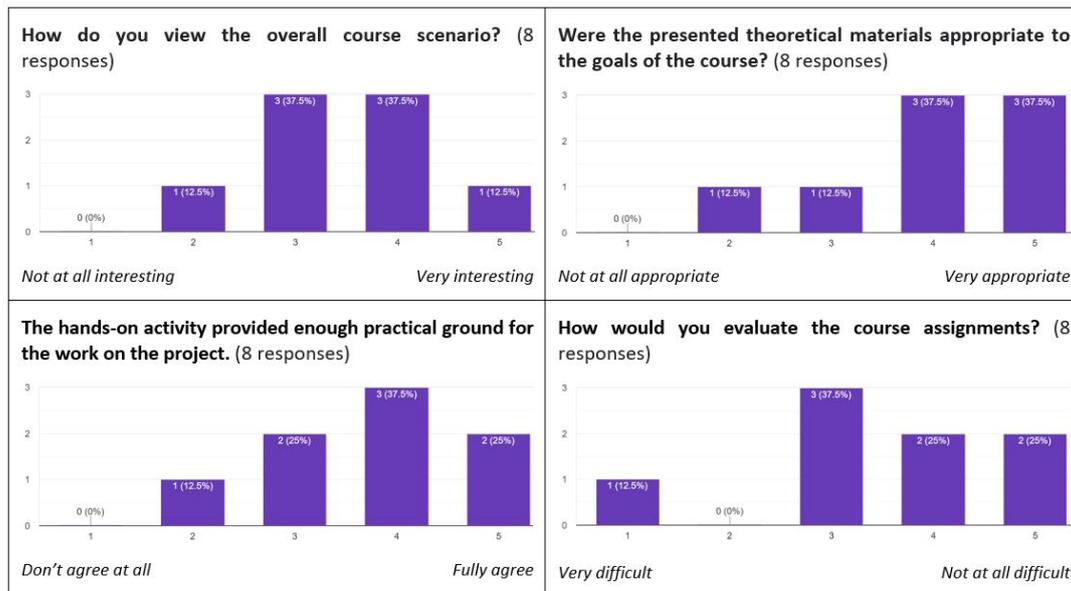


Figure 2: Evaluation of the course scenario (course questionnaire).

Further details were obtained via open questions asking the students to enumerate some of the strong and weak points in the course approach and provide suggestions for improvement and general comments about whether or not the course was appropriate for the BCE programme, its potential (or lack thereof) from a pedagogical perspective, whether or not it was useful as compared with their other projects, etc. Strong points mentioned included being introduced to “a lot of tools that can be used”, the combination of “practical work together with theoretical work” that enabled them to “learn better” and the fact that the course was “practical, descriptive” and covered “wide areas of interest”. Weak points included aspects such as difficulty in following the course “without prior knowledge of informatics”, “explanations were good but can be improved”, “too much in too little time, too little time to practice during the semester”. Additional comments referred to the relevance of the course for the BCE programme, especially from a “research point of view”, and to the fact that it taught students to “objectively look at all data and not believe something blindly”. It was suggested as well that the course should focus on fewer tools but “explain them more in depth”. Similar remarks were collected from the official evaluation. On the one hand, the students assessed that the course helped them to “gain knowledge in programming tools” and get “insight into the possibilities of computer science in the humanities”; there was a recognition that “students should be trained more in digital tools” and an appreciation that the “introduction to the different programs” can subsequently be used for “research in text interpretations”. On the other hand, it was considered that “at some points it was hard to follow the lecturer” without being a “computer programmer” and there was a sense of feeling “a bit lost with all the information”; it was suggested that the course should provide “less content but longer explanation of the most important” items and should “spend more time on a program”.

While the general idea of providing an overview of a variety of tools and methods for computational text analysis and the combination of theoretical and practical approaches seemed to have been positively perceived by the students, the project evaluations and the collected feedback suggested that the number of programs studied in

class and the time allocated to some of them should be adjusted. For this reason, during the second iteration of the course for the 2020-2021 winter semester, some tools were presented succinctly (e.g. Google Books Ngram Viewer and Voyant), while more time was allocated to explanations and sessions implying a higher degree of difficulty (e.g. the TXM specificities model, topic modelling and programming using R and Python). Another point that required further attention was related to the need for examples of how the selected methods, tools and data can be used to derive interpretations and encourage critical thinking. As Goldstone (2019, p. 2) observes when referring to literary studies, more attention should be paid in the humanities to methodologies for analysing and making a “convincing scholarly argument using quantitative evidence”. The hands-on activities in the second iteration therefore included additional discussions on the evaluation of the applied methods and the interpretation of results, as well as potential bias and limitations. Moreover, the assignment and grading descriptions contained explicit examples and criteria for formulating research questions and hypotheses according to the data intended for analysis, and for proposing interpretative scenarios with the goal of fostering self-reflection on the experience as a whole (Figure 1, dotted output). The project reports from this second round showed increased awareness in articulating research questions and interpretation arguments based on the analysed data. These aspects will be further considered and developed in the next iteration of the course during the 2021-2022 winter semester.

4 Conclusion and future work

The paper describes a digital pedagogy and evaluation framework applied to a case study combining different approaches – theoretical, application-oriented and project-based –, underpinned by a view of the DH classroom metaphorically defined as a “node”. The metaphor encompasses the hybrid character of digital humanities teaching, which involves the use of various categories of study materials, themes, examples, methodologies and skills, as well as the idea of exchange and connection with other disciplines. It was assumed that the application of this type of hybrid approach in the course design, together with the coverage of a variety of tools to be studied in class, might foster a reflective mindset, enabling the students to contextualise the digital technologies within different humanities areas, beyond the acquisition of technical skills and knowledge of computational methods and tools usually referred to as the toolbox-based education paradigm.

The case study involved a small number of students and two iterations of the teaching experiment (the first being described in more detail in the paper). Although the findings are currently limited and require further investigation, an evaluation of the students’ projects and analysis of their feedback seem to suggest that a combination of the toolbox- and mindset-based pedagogical perspectives (or a setting in which the second includes the first) would allow the DH classroom to be framed as a node within the humanities curriculum network. The main “entry points” of such a node pertain to hands-on experience with data selection, preparation and assessment and with tools for analysing the selected data, basic theoretical insights into the computational methods used for analysis beyond the black-box appearance of the interface, contextualisation of the analysis within a thematic, meaningful framework, interpretation examples including potential pitfalls, bias and limitations, and references to other disciplines as fields of application and further reading of related humanities and technical literature. The “exit points” of the node refer to a set of acquired skills in

computational analysis, critical thinking and self-reflection that may serve in formulating research questions and hypotheses, interpreting results and reflecting on the implied experience, exploring new paths and making connections, and understanding technology as both a tool and an asset for continuous questioning and discovery within a humanistic area of enquiry. Even though such a model needs refinement and more iterations of the suggested course scenario, together with experimentation with other syllabi for teaching digital technologies in the humanities, it proposes a framework that may be assessed, validated or invalidated, modified and improved, and possibly an avenue for further research and debate on ways and metaphors of imagining the shift from a toolbox- to a broader mindset-based paradigm in digital humanities pedagogy.

Acknowledgement

The author would like to thank Sarah Cooper, from the Language Centre of the University of Luxembourg, for English proofreading.

References

Anthony, L.

2019. AntConc. (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>.

Bourdieu, P.

2005. *Language and Symbolic Power*. Polity Press, Cambridge, UK, Malden, US.

Cordell, R.

2019. Teaching Humanistic Data Analysis. Talk delivered on 2 May 2019 as part of Gale's Digital Humanities Day at the British Library. <https://ryancordell.org/research/teachingHDA/>.

Cunningham, H., V. Tablan, A. Roberts, and K. Bontcheva

2013. GATE. [Computer Software]. <https://gate.ac.uk/>. Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. *PLoS Comput Biol* 9(2): e1002854. doi:10.1371/journal.pcbi.1002854. <http://tinyurl.com/gate-life-sci/>.

De Marneffe, M.-C. and C. Manning

2008. Stanford typed dependencies manual. September 2008, revised for the Stanford Parser v. 3.7.0 in September 2016. https://nlp.stanford.edu/software/dependencies_manual.pdf.

Fišer, D., J. Lenardič, and T. Erjavec

2018. CLARIN Resource Families. [Data Collection]. <https://www.clarin.eu/resource-families>. CLARIN's Key Resource Families. In *LREC 2018*. <https://www.clarin.eu/sites/default/files/LREC2018-families-final.pdf>.

Goldstone, A.

2019. Teaching quantitative methods: What makes it hard (in literary studies). In *Debates in the Digital Humanities 2019*, M. Gold and L. Klein, eds. University of Minnesota Press. <https://doi.org/10.5749/j.ctvg251hk>.

- Hajlaoui, N., D. Kolovratnik, J. Vaeyrynen, R. Steinberger, and D. Varga
2014. Digital Corpus of the European Parliament (DCEP). [Data Collection]. <https://ec.europa.eu/jrc/en/language-technologies/dcep>. DCEP-Digital Corpus of the European Parliament. *Proc. LREC 2014 (Language Resources and Evaluation Conference)*, pp. 3164-3171. Reykjavik, Iceland, May 26-31, 2014. http://www.lrec-conf.org/proceedings/lrec2014/pdf/943_Paper.pdf.
- Hart, M.
2004. Project Gutenberg. [Data Collection]. <https://www.gutenberg.org/>. The Project Gutenberg Mission Statement. https://www.gutenberg.org/about/background/mission_statement.html.
- Hazard, P.
1961. *La crise de la conscience européenne, 1680-1715*. Fayard, Paris.
- Heiden, S., J.-P. Magué, and B. Pincemin
2010. TXM. [Computer Software]. <http://textometrie.ens-lyon.fr>. The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. In *24th Pacific Asia Conference on Language, Information and Computation - PACLIC24*, R. Ootoguro, K. Ishikawa, H. Umemoto, K. Yoshimoto, Y. Harada, eds., pp. 389-398. Institute for Digital Enhancement of Cognitive Development, Waseda University, Sendai, Japan. <https://halshs.archives-ouvertes.fr/halshs-00549764>.
- Jockers, M.
2015-2017. Syuzhet. [Computer Software]. <https://github.com/mjockers/syuzhet>.
- Klein, D. and C. Manning
2013. Stanford parser. [Computer Software]. <https://nlp.stanford.edu/software/lex-parser.shtml>. Accurate Unlexicalized Parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 423-430. <https://nlp.stanford.edu/~manning/papers/unlexicalized-parsing.pdf>. See also De Marneffe and Manning (2008).
- Kripke, S.
1981. *Naming and Necessity*. Blackwell Publishing, Malden, MA, US, Oxford, UK, Carlton, Victoria, Australia.
- Liu, A.
2013-2017. Digital Humanities Resources for Project Building — Data Collections & Datasets. [Data Collection]. <http://dhresourcesforprojectbuilding.pbworks.com/w/page/69244469/Data%20Collections%20and%20Datasets>.
- Lynch, K.
1997. *The Image of the City*. The MIT Press, Cambridge, Massachusetts, and London, England.
- Mahony, S. and E. Pierazzo
2012. Teaching Skills or Teaching Methodology? In *Digital Humanities Pedagogy: Practices, Principles and Politics*, B. D. Hirsch, ed. Open Book Publishers. <http://books.openedition.org/obp/1639>.

- McCallum, A.
2002. MALLET: A Machine Learning for Language Toolkit. [Computer Software]. <http://mallet.cs.umass.edu/>.
- Michel, J.-B., Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, et al.
2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331(6014):176–182. <https://science.sciencemag.org/content/331/6014/176>.
- Moretti, F.
2013. *Distant reading*. Verso, London, New York.
- Papadopoulos, C. and S. Schreibman
2019. Problem/Project Based Learning in the DH Classroom: Situated Learning, Empowerment, and Knowledge Creation. In *DH Benelux Conference, 11-13 September 2019*. University of Liège, Belgium.
- Pennebaker, J. W.
2013. *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury Press, New York, London, Oxford, New Delhi, Sydney.
- Řehůřek, R. and P. Sojka
2010. gensim – Topic Modelling in Python. [Computer Software]. <https://github.com/RaRe-Technologies/gensim>. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50. ELRA, Valletta, Malta. <https://is.muni.cz/publication/884893/en>.
- Rinker, T. W.
2015-2018. Sentimentr. [Computer Software]. <https://github.com/trinker/sentimentr>.
- Rockwell, G. and S. Sinclair
2016. *Hermeneutica: Computer-Assisted Interpretation in the Humanities*. The MIT Press, Cambridge, Massachusetts, London, England.
- Ross, J.
2017. Speculative method in digital education research. *Learning, Media and Technology*, 42(2). <https://www.tandfonline.com/doi/full/10.1080/17439884.2016.1160927>.
- Russell, J. and M. Hensley
2017. Beyond buttonology. Digital humanities, digital pedagogy, and the ACRL Framework. *College & Research Libraries News*, 78(11). <https://crln.acrl.org/index.php/crlnews/article/view/16833/18427>.
- Sinclair, S. and G. Rockwell
2012. Teaching Computer-Assisted Text Analysis: Approaches to Learning New Methodologies. In *Digital Humanities Pedagogy: Practices, Principles and Politics*, B. D. Hirsch, ed. Open Book Publishers. <http://books.openedition.org/obp/1644>.
- Sinclair, S. and G. Rockwell
2016. Voyant Tools. [Computer Software]. <https://voyant-tools.org/>.

- The Google Ngram Viewer Team
2011. Google Books Ngram Viewer. [Computer Software]. <https://books.google.com/ngrams>. See also Michel et al. (2011).
- Thompson, J. B.
2007. *Ideology and Modern Culture: Critical Social Theory in the Era of Mass Communication*. Polity Press, Malden, US.
- Thompson, P.
2004. Black Immigrants to Britain, 1890-1975. [Data Collection]. UK Data Service, SN: 4936. <http://doi.org/10.5255/UKDA-SN-4936-1>.
- Underwood, T.
2019. *Distant Horizons. Digital Evidence and Literary Change*. The University of Chicago Press.
- Vonnegut, K.
2010. Short lecture by Kurt Vonnegut on the “simple shapes of stories”. Posted by David Comberg, 30 October 2010. <https://www.youtube.com/watch?v=oP3c1h8v2ZQ>.
- Whitby, G.
2007. Pedagogies of the 21st Century. In *ACEL 2007 International Conference*. Sydney, Australia. <https://robertsonss.eq.edu.au/Supportandresources/Formsanddocuments/Documents/pedagogy-for-the-21st-century.pdf>.