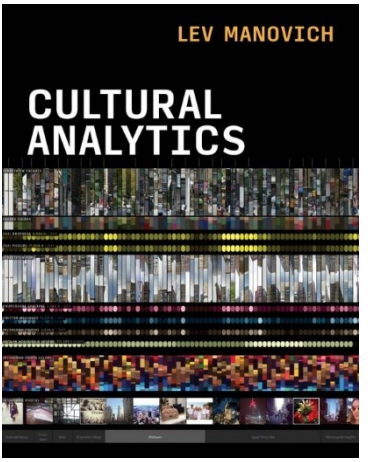
Collect, analyse, visualize data: establishing the field of Cultural Analytics in Digital Humanities

Liliana Vasques

UNIVERSITY OF COIMBRA | CLP
ORCID: 0000-0002-3519-3872



Lev Manovich, *Cultural Analytics*, Cambridge, MA: The MIT Press, 2020, 336 pp. ISBN 978-0262037105.

Lev Manovich has been a prolific scholar and theorist through the 2000's, closely researching digital culture and its many facets and developments. We can trace a coherent and reflexive production among numerous articles, projects and books. *Cultural Analytics* can be situated in a trajectory in which we can highlight *The Language of New Media* (2000) and *Software Takes Command* (2013) as staples of his work concerning digital culture and media theory. All through the monographs, we find an extremely well balanced integration of facts and arguments from fields such as computer science, media theory, art history, visual arts and software studies to address digital culture. In *The Language of New Media*, the author takes the task of identifying and explaining how the use of the computer to create media content determines the characteristics of that content. What follows,

MATLIT 9.1 (2021): 273-276. ISSN 2182-8830
DOI: 10.14195/2182-8830_9-1_16

in *Software Takes Command*, is necessarily a thorough analysis concerning the 'softwarization' of social interactions, work, art, and culture.[1]

Manovich has the ability of permanently keeping up with the evolving nature of digital culture and, as such, *Cultural Analytics* is published when big data is a fully established concept and instrument in Science, Politics, Economics, etc., and its broader proposition is to establish a field to engage with big data when it comes to culture. Throughout the chapters, the author provides a theoretical and practical framework to "systematically study global media cultures today"(17) thus laying the groundwork for the field of cultural analytics in Digital Humanities. The concept of cultural analytics took a growing importance in the context of the Software Studies Initiative (2007) laboratory and the research developed by Manovich and his students (some of these projects are available at http://lab.softwarestudies.com/). The book intends to trace this new field, strongly supported by data science, data visualization, and media theory while, simultaneously, not avoiding 'expected' questions and facing possible limitations of this approach. The titles of several chapters suggest precisely the methodological fascination with the datafication of culture but also inevitable epistemological interrogations: "Looking at culture with computers", "The Science of culture", "Representing culture as data", "The sematic gap", "Do we explain culture?". At the same time, the structure of the book as reflected in the order of chapters and subchapters seem to be in synch with the main goal of establishing the field of Cultural Analytics – theoretical Framework, methodology and results, conclusions about Cultural Analytics and its relationship with culture, art, and the Humanities. A conventional and useful structure, given its didactical nature and practical applications, which, concurrently, is traversed by a permanent debate on the role of Cultural Analytics.

The "Introduction" and "Studying Culture at Scale" chapters provide us with a global view of the field, relating it, namely, to the development of Statistics and computational methods of data analysis. Furthermore, they provide arguments for the use of Cultural Analytics and the need to rethink methodological tools of the Humanities. A perfect epitome of this book, "Looking at culture with computers" (the title of an introductory subchapter) is the answer to study contemporary culture considering the scale at which content, interactions, and events are created. According to the author, it is inevitable to use quantitative, data science methods if we want to fully understand a culture that is digital, exponential, networked, and producer of massive amounts of data. Notwithstanding, Manovich clarifies that

[1] Cf. Manuel Portela, "New Media as Software" in *MATLIT* 1.2 (2013): 176-80. https://impactum-journals.uc.pt/matlit/article/view/2182-8830_1-2_16

the book does not address data analytics and its methods (inferential statistics, experimental design, machine learning, text analysis, geospatial analysis, music data analysis, network analysis), but rather cultural analytics, i.e., how to turn culture (events, actions, media, etc.) into data. Simply put, the work is concerned with the prior step in the process of analysis. Cultural analytics is, thus, put forward to "explore massive collections that may contain billions of images and videos" and "as opposed to only historical culture" (7), allow the analysis of all kinds of media (not only texts). Manovich states, clearly, that his view is one of inclusiveness and complementarity of methods in which the use of big data to understand culture is the possibility to "augment our human abilities by providing new interfaces and techniques for observing massive cultural datasets and flows" (11). An ecology of methods, where qualitative and quantitative analysis are combined to obtain a more comprehensive understanding of our cultural present.

Chapters 2 and 3 deal with methodological aspects such as research sources, types of data, data selection, and methods for turning gathered information into data. The book provides us with case studies to better understand how to use this analytic proposition. In chapter 2, Manovich writes about two major topics – types of data and cultural sampling. Firstly, he identifies four types of data through which we can study digital culture – media, behaviours, interactions, and events. One can ask if this categorization captures all the important materials and/or contents. Where, for example, should we include texts and discourse? How can we account for 'non-digital' cultural life? Concerning cultural sampling, the author discusses the relationship between sampling (selection) methods and inclusiveness, representativeness, diversity, and biases. Analysing the use of digital archives, museums and libraries, he revisits the long existing question about who chooses what and under what criteria. Sampling is a complex task and demands a fairly good understanding of the cultural subject in hand. In this context, Manovich advocates for the use of big data to overcome the limitations of cultural sampling. Using big data instead of small samples will guarantee diversity of artefacts, people, and other data to be preserved. Furthermore, according to the author, it will also enable the field of cultural analytics to consider both patterns and individual data.

In chapter 3, he describes a new paradigm to represent data based on the same principle as the use of big data (an alternative to sampling) – avoiding reduction. Manovich calls it media visualization. He describes it as a "deeply qualitative rather than quantitative approach" (223) that uses metadata contained in images to sort and group them into categories. This method dismisses the need to translate text, images or videos into new signs such as points and lines – reduction. He seems especially interested in the

application of media visualization to images. The book includes high quality images of different research projects where data visualization was applied. Considering the content of previous chapters, Cultural Analytics seems to be more adequate to digital culture and image visualization methods. The appeal of such visualizations has been growing and, lately, these have been developed in such a manner that analytic images and graphs came to stand on their own as artistic content. Data visualization art has become a new field of its own.

The Conclusion collects ideas from the previous chapters and reiterates the importance of engaging with the present characteristics of culture in order to ask the 'right' questions and to look at it in new ways. Manovich expresses the excitement about the possibility of following global digital cultures in real time, a future direction he aims to explore. Inevitably, this raises obvious questions. Should we track, monitor and record culture in this manner? Our behaviour is now fully traceable. Can cultural analytics lead us to a certain fetishism and obsession towards the idea of tracking the totality of cultural events? All methods carry limitations and risks. The author acknowledges that fact in a fairly explicit way. "The semantic gap" and the "Do we want to explain culture" subchapters address some of the epistemological issues. One major aspect of using cultural analytics methods is how to engage in a critical inquiry of results to understand cultural events. Another challenge put forward by Manovich himself is whether we can use this approach to "question stereotypes, assumptions, concepts, and existing knowledge about culture" (248). When it comes to Digital Humanities, a field of study in which cultural analytics methods should be used, according to Manovich, arguments against the excessive quantification and scientification are well known. Humanities should not try to emulate the 'hard sciences' or the 'social sciences' and risk losing their specific purpose. Furthermore, Digital Humanities should not be about computing phenomena, but rather developing reflexive and critical thought. Manovich is aware of this and stands by an ecology of qualitative and quantitative methods in the study of culture. Nonetheless, he cannot prevent himself from sharing the 'big data dream' – a crossover between Google Analytics, super visualization, and Gapminder World Interactive "able to monitor a full range of aesthetic and semantic characteristics of cultural media and subtle but crucial differences between numerous cultural artifacts, especially at a global scale" (35).