Digital Humanities: Ethical Implications and Interdisciplinary Challenges

Maria Eunice Gonzalez
São Paulo State University (UNESP/Marília)
Department of Philosophy
Av. Hygino Muzzy Filho 737, Marília/SP - Brazil
Email: eunice.gonzalez@unesp.br

Mariana Vitti Rodrigues
São Paulo State University (UNESP/Marília)
Department of Philosophy
Av. Hygino Muzzy Filho 737, Marília/SP - Brazil
Email: mvittirodrigues@gmail.com

Abstract:

What is meant by Digital Humanities? This paper presents and discusses characterizations of Digital Humanities, with emphasis on methodological and ethical problems related to the use of Big Data in human sciences. There is still no consensus regarding the definition of Digital Humanities, but there is a general agreement that collaborative inter/multi/transdisciplinary practices are involved in the study of computational processing of massive amounts of data available in areas of human sciences. Studies concerning digital resources, electronic publishing, and digitized historical newspaper collections, among others, have played a prominent role in these areas. This paper addresses questions about the genuine interdisciplinary nature of research on Digital Humanities, and discusses ethical issues associated with Big Data analytics, from a complex systems perspective. Emphasis is given to implications (positive and negative) of the use of information and communication technology in studies of the dynamics of human autonomy. To conclude, we discuss possible upcoming challenges of the extensive use of Big Data technology for the development of Digital Humanities in the domain of human autonomy.

Keywords: Digital Humanities, Big Data analytics, meaning, interdisciplinarity, human autonomy

Introduction

The notion of Digital Humanities (henceforth DH) arose around the year 2000, as a development of the project "Computation in Humanities" that started in the 1950s. Although there is no consensus regarding the definition of DH, there is a general agreement that it involves collaborative inter/multi/transdisciplinary practices in the generation and analysis of digital data and computational processing available in the disciplines of the humanities.

In general terms, DH emerged with the interplay between traditional humanities and computational methods of designing, computing, and modelling human activities in a variety of

domains (Burdick et al., 2013). The extensive use of algorithms, standardization of data for packing and analysis, and the possibilities for reuse, reanalysis, repurposing, and recombination, among others (Leonelli, 2014; Collmann et al., 2016), have challenged the standard methods in humanities. In addition, DH encompasses the digitalization of knowledge and historical artefacts, involving structuring and modelling, including the analysis of data from social media.

The development of Artificial Intelligence and Big Data Analytics in the 21st century, along with Information and Communication Technologies (ICTs), has empowered DH, altering practices and forms of knowledge dissemination. Although there is no precise definition of the term Big Data (BD), it may be understood as describing the technology of mining, cleaning, modelling, storing, and analysing massive amounts of a wide variety of data, generated at high speed, with current computational power. In general, the data are not well structured and have to be integrated in databases, in accordance with specific objectives, in order to acquire epistemological, political, or economic value. In this technological scenario, Big Data was initially characterized in terms of the 3 V's: velocity, volume, and variety (Laney, 2001).

As a consequence of the extensive use of Big Data and statistical tools, scientific practice has faced a growing process of automation, especially in the elaboration of probabilistic/algorithmic models. Inspired by Breiman (2001), Pietsch (2013) argues that Big Data analytics allow a shift in modelling from explanatory hierarchical modelling, which encompasses few variables to account for a wide range of the objects studied, towards predictive algorithmic modelling that changes according to the database.

The humanities were not left out from the overflow of new computational technologies, and new modelling assumptions, imported from the formal sciences, are growing in DH. In linguistics, for example, by applying complex algorithms for natural language processing, the predictive power of machine translation has increased according to the input received from users. Automated translation does not require a deep understanding of the general rules of the structure of language, as advocated by Chomsky. Instead, it requires massive data acquired from users' inputs, together with probabilistic algorithms. A well-cited example is the statement of Frederick Jelinek, a former IBM researcher, concerning speech recognition: "Every time I fire a linguist, the performance of the speech recognizer goes up"[1].

Areas of social sciences have also incorporated the development of computational tools that allow the investigation of social phenomena, which brings new possibilities of communication, but also new challenges. For instance, with the massive amount of available data, models of the dynamics of public opinion have emerged, aiming at understanding collective opinion. In the realm of arts, methods to analyse and produce artistic works have emerged in unprecedented ways, with multimedia installations and automated analyses leading to new possibilities for the creation as well as evaluation of artistic material.

Philosophy has not escaped from this revolution in DH. The developing field of computational philosophy aims at analysing problems related to philosophy of science,

philosophy of language, social epistemology, and philosophy of mind, among others, by applying computational techniques (Grim & Singer, 2020). A novelty in this area is that computational philosophy is not the same as philosophy of computers or philosophy of technology; what characterizes this emerging field is the use of computational techniques for the study of philosophical problems.

The present paper addresses questions about the genuine interdisciplinary nature of research concerning DH, as well as ethical issues associated with Big Data analytics, from a complex systems perspective. Focusing on semantic and pragmatic aspects of automated human activities, the following questions are addressed: (i) How could genuine interdisciplinary research be developed in Digital Humanities? (ii) What is the role of meaningful information in the digitalization of human events? Our working hypothesis is that research in human sciences requires understanding of the complex dynamics of events where a relational concept of meaning plays a fundamental role.

Questions (i) and (ii) will be investigated in two steps. In the second section, a characterization of Digital Humanities is proposed, inspired by Complex Systems thinking. In the third section, the notions of data and meaning are introduced, in the context of the growing datification of DH. In addition, an example of a mechanical generator of meaning is presented, in the domain of natural language, as indicative of the role of new methodological practices in DH. Finally, ethical challenges of DH are considered for the maintenance of autonomy of researchers, arguing for a strong interdisciplinary commitment to the development of ethically sound research in humanities.

Digital Humanities: A perspective from complex systems thinking

> The wars over the core in the humanities have contributed to a malaise in which the humanities are widely perceived as "irrelevant," lacking the practicality of business, law, or medicine.
> (Burdick et al. 2013, 16)

Since the coining of the term 'Digital Humanities', attempts to define it have flourished. Callaway et al. (2020) built up a corpus with 334 attempts to characterize DH, in which they identified four 'clouds' of words, associated with four general topics mechanically extracted from the analysed papers by topic modelling algorithms. These four topics were given names considered representative of each set: (i) Code, (ii) Community, (iii) Distant Reading, and (iv) Diversity and Inclusion. A curious commitment, suggested by Callaway et al. (2020), is that the validation of arguments, involving uses of quantitative and qualitative methodologies from diverse disciplines, requires that the interpretive work of a specific type of knowledge should be described in terms of the connection of the topic to a specific DH area. In the end, conclusions based on a quantitative analysis of text documents indicated the obvious, namely the polysemic use of the term DH concerning the issue of interdisciplinarity.

By analyzing the use of the term DH in different scientific communities, Roth (2019) proposed that the notion of Digital Humanities could be understood considering three different aspects, regarding: (a) research undertaken in contemporary humanities, encompassing the community whose work is performed using digital tools such as digital archives; (b) the use of quantitative methods in order to develop abstract models to understand human phenomena, such as agent-based modeling in social studies; (c) the study of human-computer interaction and the analysis of online communities.

Interpreting the Roth classification, Piotrowski (2020) suggested that the first sense of the term DH, which he called *Contemporary Humanities*, refers to humanists that simply incorporate digital tools in performing their habitual practices with their traditional methods. According to the author, this does not involve interdisciplinary commitment, with every discipline having its own home. In contrast, the second sense of DH, which he called *Computational Humanities*, relates to the use of new methodologies, along with the creation and manipulation of models by computers:

> And this is exactly where the root of the problem lies: the majority of actors in "the field" (to use a maximally neutral term) belong to the "contemporary" or "digitized" humanities, which can be seen as a natural and organic development *within* the existing disciplines. The actors in the "numerical" or - as I am going to argue: "computational" - humanities, however, are not only a minority, but they are typically the ones with an explicit DH denomination or affiliation. (Piotrowski 2019, 4)

Piotrowski argues that DH should be equated to *Computational Humanities,* which is strictly composed by the community that applies computational models in a given domain and also creates abstract models, such as metamodels potentially manipulated by computers. He concludes that *Computational Humanities* should be considered a new discipline that gathers knowledge from computer science and the humanities. From this perspective, DH would have a disciplinary character, despite being constituted by different methods imported from other disciplines.

In a different vein, Burdick et al. (2013), instead of defining DH as a discipline, proposed that, generally speaking, DH should be characterized by "[...] the opportunities and challenges that arise from the conjunction of the term *digital* with the term *humanities*". We understand, nevertheless, that the challenges should be faced according to an interdisciplinary approach, taking into consideration methodological and ethical aspects. To do so, we propose a characterization of DH from a Complex Thinking perspective, by inquiring: How could a genuine interdisciplinary research practice be developed in Digital Humanities?

At the present stage of development, the following methodological scenario dominates sciences in general: On the one hand, in formal sciences, mainly quantitative methods of investigation are applied; scientists undertake computational methodologies and, more importantly, focus on kinds of objects that can be easily formalized by the process of standardization. On the other hand, in human sciences, qualitative methods are mostly applied, which strongly rely on

interpretative assumptions that are historically and contextually embedded[2]. The challenges faced by the so-called digital humanists are altering the ways that researchers act in academia, arts, literature, history, and social sciences, among other disciplines. However, it is hard to see which changes have been undertaken in the hard/formal sciences under the influence of humanities.

Far from advocating a methodological unity of science, as was the aim of the logical positivists of the 20th century, our proposal here is an intercultural, honest, and critical dialogue among researchers from different sciences, recognizing the incompleteness of the various types of knowledge. We believe that DH should be developed as a field of multi/inter/transdisciplinary research grounded on critical and genuine dialogue that considers epistemological and ethical implications of uses of digital technology in the study of humanities.

Challenges brought about by DH should be characterized as problems of organized complexity that, according to Weaver (1948, 539), "[...] are all problems which involve dealing simultaneously with a *sizable number of factors which are interrelated into an organic whole*". To tackle such problems, in addition to the use of new technologies of Big Data within the sciences of complexity, which develop sophisticated techniques for modelling, there is the need to focus on perspectives of complex systems thinking.

Inspired by Zoya & Aguirre (2011), we propose a characterization of DH that considers the study of the problems of organized complexity that involve technical aspects, as well as ethical and political issues. In this way, there would be complementarity between complex thinking and complex (hard and soft) sciences:

> Complex thinking can provide the reflective field necessary to develop an inclusive epistemic framework of ethical and political values consistent with the needs and challenges of social, local, national, regional, and planetary communities. The sciences of complexity can provide concrete methodological tools for the study of complex phenomena. We believe that the most fundamental challenge is to stimulate the development of the sciences of complexity guided by complex thinking. (Zoya & Aguirre 2011, our translation)

Zoya and Aguirre indicate a way to deal with our initial question, (i): How could a genuine interdisciplinary research practice be developed in Digital Humanities?

Our suggestion is that research in DH, involving statistics, data science, computer science, data visualization, and ontology, among others, should be developed according to an approach that considers, as far as possible, the multiple dimensions of the studied object, together with its epistemological and ethical implications. In short, investigations based on complex systems thinking should include the following presuppositions:

i) Multilayer analysis of events, ranging from micro to macro levels;
ii) Attention towards phenomena emergent from the interplay of different disciplines;

iii) The relevance of circular causality in DH practices that might be affected and affect other areas, modifying the object of investigation.

Presuppositions (i)-(iii) require a collaborative perspective to be developed in DH that comprises qualitative and quantitative approaches to express the complexities underlying human, environmental, and political events. Thus, Digital Humanities could be characterized as a genuine collaborative dialogue among social, formal, and hard sciences, by emphasizing shared techniques for modelling, analysing, and interpreting data. DH would then be a genuine multi/inter/transdisciplinary area of study, providing a multi-way road towards scientific research in which the main aim should be to benefit society.

A core challenge in contemporary DH that contemplates the proposed collaborative perspective is the attempt to deal with topics, investigated in the next section, about the semantic and pragmatic aspects of human activities involving the application of Big Data analytics resources, in accordance with a Complex Systems perspective.

Meaning as the core of humanities

The digitalization of information concerning human events, in order to express them as units of data that can be structured in integrated databases for mechanical analysis, is a hard task that involves qualitative and quantitative criteria of relevance. To begin with, the study of human events requires a careful approach that takes into consideration multiple dimensions of analysis, such as historical and contextual foundations. For instance, analysing a human trait such as dishonesty by applying only quantitative methods that search for metadata and correlational patterns among variables may lead to disastrous results[3]. To illustrate this claim, we discuss the notion of data and the process of meaning generation, indicating challenges of an automated approach to meaning.

Some methods imported by DH from formal sciences mirror the presupposition that data are a commodity that shares quantitative, meaningless aspects of the reality, that the so-called 'raw data' can be digitalized without much loss. In this regard, Drucker (2013, 2) alerts: "[...] the digital humanities can no longer afford to take its tools and methods from disciplines whose fundamental epistemological assumptions are at odds with humanistic methods". The author advocates that the idea of raw data as a mirror of reality, providing a faithful reflection of such reality, should be revisited by data-driven research in humanities.

The revision of what counts as data in humanities should not open the doors for relativism. It is not to say that data has no connection to reality, but that this connection is a temporary framework communicated through sign-relations, as partial expressions of the object under scrutiny.

The concept of raw data has been investigated by Leonelli (2015, 813), among others, who calls the attention to the following paradox, which: "[...] consists of the observation that, despite their epistemic value as 'given', data are clearly made." She suggests a relational concept of data

that are intrinsically context-dependent, with their epistemic value depending on the context in which the data are collected and stored, and the context in which they are analysed.

In order to apply computational techniques to vast amounts of data stored and integrated in massive databases, the data have to be ready for *data journeys* (Leonelli, 2014). Data journeys involve de-contextualizations, re-contextualizations, and, most interesting, the possibility of reuse of data. Adding to that, Collmann et al. (2016) emphasized the possibility for reanalysis, recombination, and repurposing of data. They indicated that when data are *generated*, there is usually a purpose for which they are *meant*. However, reusing, reanalysing, recombining, and repurposing the data may change their primary meaning. The context-dependency of data may change their epistemic values, according to their journeys, with ethical and epistemological implications for science and society.

In short, once integrated in structured or semi-structured databases, data can grow in meaning by the process of re-contextualization, which could re-direct their original purpose when generated in a given context. According to our working hypothesis, research in human sciences requires understanding of events where a relational concept of meaning plays a fundamental role. Our challenge then is to characterize meaning in such a way that highlights its relational aspects in the context of DH.

In the history of philosophy, one can find many attempts to characterize the notion of meaning. A common hypothesis, which interests us here, is that meaning is an emergent property of the relationship between the dispositional patterns of a given environment and the perception/action of embodied and embedded organisms.

Highlighting the perspective of the organism, von Uexküll (1982 [1940], 2010 [1934]) proposes the notion of meaning in terms of the 'species-specific environment' to express the recurrent cycle of perception and action in which organisms are shaped by the physical constraints of the specie. (Uexküll 2010 [1934]). The author exemplifies his hypothesis in the following passage:

> If we stand before a meadow covered with flowers, full of buzzing bees, fluttering butterflies, darting dragonflies, grasshoppers jumping over blades of grass, mice scurrying, and snails crawling about, we would be inclined to ask ourselves the unintended question: Does the meadow present the same view to the eyes of so many various animals as it does to ours? (Von Uexküll, 1982 [1940], 45)

Von Uexküll argues that the constraints of our species may indicate different possibilities for us to explore, contemplate, and enjoy the meadow, while the meadow offers dispositions for action for other beings that have different constraints.

In a view not epistemologically distant, but methodologically distinct from that of Von Uexküll, Gibson (1956, 1979) emphasizes the role of niches, characterized as those parts of the environment containing engravings of the relationship between the environment and the perception/action of different species. In his *Ecological Philosophy*, the concept of *affordance* is

proposed to represent meaningful information available for embodied and embedded agents, allowing the occurrence of actions in a direct way that is not mediated by subjective representations. Thus, the conceptions of meaning proposed by Gibson and Von Uexküll are relational, in that they emerge from the dynamic agent-environment relationship. However, they differ methodologically, since Gibson asserts that meaning is not dependent upon the subjective universe of the animal, but instead is an emergent property of the dynamic relation between the action of the organism and its natural niche.

In the context of digital objects, *technological affordances* can be understood as types of emergent second order dispositions embedded in the context of digital niches. Such second order dispositions allow for investigations of meaningful perception/action in virtual reality. One example of a technological affordance is the use of GPS, for motorists familiar with this kind of gadget. In the case of DH, indications by recommender systems (RS) of possible relevant papers that might drive people to further reading, in a specific context, provides another example of technological affordances. A common property of these examples is that they both presuppose an interactive familiarity with the digital environment, in a way that allows the user to navigate appropriately and perform meaningful actions.

The relational aspect of meaning is also highlighted by Peirce (CP 5.402), who emphasises the role of signs in the establishment of the agent-environment relationship. In his theory of signs (or semiotics), he claims that the meaning of a sign (characterized as something that stands in the place of something else[4]) lies in the possible consequences that the perception of a sign can play in the action of an agent in a given context (Peirce, CP 5.402). An interesting aspect of the approach of Peirce is the presupposition that signs grow in meaning through the process of thinking and doing. It is by working with signs that one can exert self-control, to adjust the held beliefs in accordance with the environment.

To sum up, the three approaches to meaning outlined here emphasise the relational property of the agent-environment relationship. Assuming that meaning is an emergent property of perceived informational patterns (which indicate law-like relationships) and actions of embedded and embodied organisms in virtual and/or ecological environments, one might ask: Can a machine generate meaning in the relational sense?

Inspired by a biosemiotic notion of meaning, grounded on the approaches of Uexküll and Peirce, Menant (2013) provides a positive answer to the above question, proposing a mechanistic Meaning Generator System (MGS hereafter). His main hypothesis is that humans, machines, and non-human animals have constraints (general and specific) that allow them to generate meaning when faced with information from the environment. According to Menant, meaning is relational and encompasses three elements: (i) the constraints of the system, (ii) the incident information from the environment, and (iii) the emergence of meaning from the relationship between (i) and (ii). Menant argues that while human and other animals generally have constraints on survival acquired during their evolutionary history, machines, in contrast, depend

(so far) on designers to specify their constraints. In both cases, he claims that meaning is a relational property that emerges from the interaction between agents (including specific constraints) and the environment (in a given context).

In short, according to the MGS, a given system encompasses a meaning generator subsystem whose role is to compare the information received from the environment to the established or learned constraints. The flexibility and origin of the constraints of the system depend on the kind of system analysed. Thus, the MGS can generate meaningful information that guides the action of the system, which, in turn, might modify its surroundings, generating new affordances that are sometimes perceived by other systems that, in a circular feedback, may grasp informational patterns and generate new meanings.

A simplified version of the MGS proposed by Menant is illustrated in *Figure 1*.

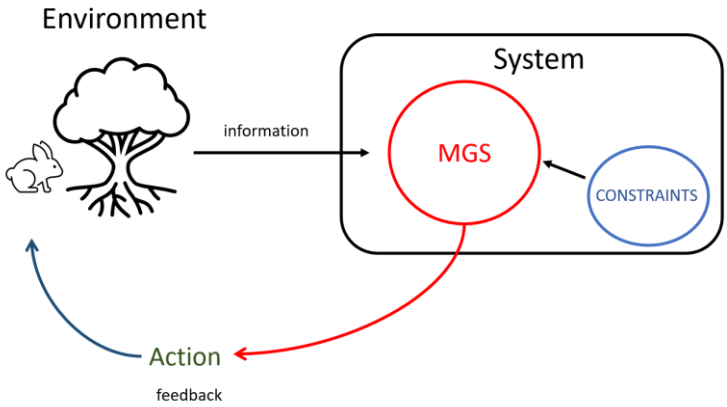

Figure 1. A simplified version of the MGS (adapted from Menant, 2009).

With the growing development of deep learning algorithms, the possibility for artificial systems to adjust their constraints to new input data may shed light on the future of mechanical meaning generation. This suggests the need for an ethical reflection on the growing possibility of the automation of meaning generation within humanities. In the next section, possible ethical consequences of the growing automation of human practices are discussed in relation to autonomous action.

Autonomy and the mechanical generation of meaning in the Big Data era

To what extent could mechanical generation of meaning, incremented by Big Data analytics, modify well-established practices of human autonomous action? As indicated, the extensive use of computers in DH has led to growing automation of scientific practices in the human and social sciences. In the Big Data era, automated analyses of data require structuring of data that, once integrated in massive databases, might provide fruitful insights by means of extraordinarily powerful computer analysis. However, in the context of humanities, the retrieval

of fruitful insights from the mechanical generation of meaning requires understanding of the subject matter, which implies the need for sound and genuine interdisciplinary research.

In the area of Computational Literary Studies, for example, Da (2019, 623) highlights 'the stop word dilemma': "Keep them [the stop words] and they produce the only statistical significance you have; remove them and you have no real results". In natural language processing, 'stop words' are highly frequent words in a text (such as the words 'the', 'a', 'an', or prepositions). For the extraction of interesting patterns from a corpus of written texts, it is necessary to 'clean' the data, which usually includes the removal of stop words. However, in literary studies, small changes in the corpus or in the parameters of statistical analysis can make a relevant difference. For instance, in the context of topic modelling, Da (2019, 625) alerts that "[...] you need only to tweak small details to discover completely different topics".[5]

On the one hand, using statistical analysis for automated reading of texts, for example, allows the results to become reproducible. On the other hand, application of the same code to the same corpus can lead to the discovery of different results (Da 2019). This happens because the meaning extracted from the data is sensitive to the decisions one makes.

A relevant novelty is that the development of mechanical text analysis and the use of natural language processing in DH go beyond the gates of universities. In dealing with recommendation systems and commercial chatbots, for instance, high accuracy and explanatory power in text analysis and text generation are not always required. The growing use of language generation models outside universities indicates that DH, characterized here from a Complex Systems perspective as a set of methodological and ethical challenges, should be considered as a systemic set of activities that surpasses academic disciplinary boundaries.

Nowadays, it is common when writing an email to receive suggestions on which word should come next. The technique applied for such recommendations is also found in the translation machine: a probabilistic analysis based on a massive database, which predicts the most common word following the words already typed. The advent of powerful generative language models took this recommendation process to the next level. In the case of GPT-3, a pre-trained, auto-regressive language model that performs tasks such as translation, text generation, and arithmetic, among others, it is possible to generate full texts (journalistic, comedy, and literary, among others) with a single input. Sooner or later, generative models of natural language will account for a large part of the activity of producing texts, making it difficult to distinguish the authorship of the text (Floridi & Chiaritti 2020).

What makes generative language models powerful is the massive database used in the pre-training phase[6]. Databases comprising texts produced by humans are the building blocks for training in the language generator, so the output of this new tool could reflect human shared opinion (at least, the opinion of a given class, namely people who write in English and have access to the internet and social media). Recognizing this, Prabhu (in an interview to Strickland in 2021) pointed out that GPT-3 is "[...] not coming out of a vacuum. It's holding up a mirror".

One of the problems faced by the developers of automated language generators is that the quality of the output, whether offensive (or not), appropriate (or not), or insensitive (or not), depends on the context within which it is generated. In other words, the success of the human-machine interaction depends on the meaningful information that emerges from this kind of interaction. Some users may enjoy offensive message exchange (when playing a game, for instance), while users of a mental health chatbot may need more careful evaluation of the effects of messages. Similarly, in the area of visual classification, Misra et al. (2016, 2930) call attention to the problem of biases associated with learning algorithms trained with the use of readily-available resources that gather data previously posted by internet users. So-called *reporting bias* emerges because when users classify an image, they choose specific properties to be highlighted, to the detriment of others, according to their criteria of relevance.

Due to the rapid advance of ICTs supported by machine learning and Big Data analytics, it is difficult to be aware of the possible consequences that their use may have in the future. However, by considering our autonomy at individual and collective levels, possible positive and negative consequences might be anticipated. A characterization of *autonomy* was proposed by Gonzalez (2017), inspired by a pragmatic notion of embodied and embedded meaningful information, according to which a person P will have a degree of autonomy if at least two conditions are satisfied:

(1) P possesses *reliable* anticipatory meaningful information of choice opportunities for future actions.
(2) P has the capacity to choose, regardless of coercive external pressure.

If (1) and (2) are satisfied, autonomy can be characterized as the human rational ability to establish independent decisions, free of external imposed determinants, in such a way that the action of the individual may be self-governed on the basis of anticipatory *reliable* meaningful information. At the collective level, autonomy might express the capacity for people to act according to their own self-organizing habits that maximize common interests. Thus, the degree of collective autonomy also depends on mutual self-governed agreement based on anticipatory *reliable* meaningful information.

The requirement for reliability of meaningful information available in the digital environment in the Big Data era leads to the following difficulty faced by Digital Humanities: Given that Big Data analytics might involve reusing, repurposing, recombining, and reanalysis of the collected data (Collmann et al. 2016), how reliable is the information available for researchers and other citizens in their exercise of autonomous scientific practice?

In the context of data-driven humanities, the use of Big Data tools and models is not neutral, so the requirement for reliability necessitates ethical and epistemological criteria to ensure that some degree of autonomy might be preserved in individual and collective action. Thus, the use of quantitative and qualitative methods requires genuine interdisciplinary research

practice, where investigators have to trust each other, in order to be able to rely on the expertise of their colleagues and, by doing so, achieve a collective level of autonomy.

In the epistemological context, Alvarado & Humphreys (2017, 373) suggested the need for a new epistemology to deal with 'thick mediation', pointing out that: "[i]t is thick in the sense that the goal of the enabling channel is not to get out of the way, but precisely to get in the way and fundamentally reshape the message" They contrasted *thick* and *thin* mediation, by referring to technologies that faithfully communicate a message from a given source to a suitable receiver. In the case of thick mediation, the information communicated is stored, along with its metadata, to be available for future analysis, recombination, reuse, and repurposing. The advent of thick mediation, they claim, gives new meaning to McLuhan's sentence 'the medium is the message' because "[...] it is not simply that the medium has its own message. It is that the message itself is augmented and shaped by the medium" (Alvarado & Humphreys 2017, 373). It seems to us that McLuhan (1964, 2) had foreseen the potential of ICTs, when he explained: "[...] 'the medium is the message' because it is the medium that shapes and controls the scale and form of human association and action". But, if that is the case, would there be a place for autonomous action in the Big Data era?

The process of *datification* and the possibility of *mechanical generation of meaning* bring to the fore not only epistemological challenges faced by data-driven humanists, but also ethical dilemmas intrinsically related to data journeys. As pointed out by Collmann et al. (2016, 142), the possibility of *reusing* the data "[...] raises the responsibilities that investigators have for what happens to data once they become available to secondary investigation". The process of *repurposing* of data "[...] poses the question about the legitimacy of analysing data acquired under one privacy context and employing it in a different privacy context." Furthermore, they alert that "[...] *recombining* data poses the question about the possibility of obtaining new information not available from the constituent data set. From a privacy perspective [...] it potentially enables re-identification of individuals from data that contains no identifier."

To sum up, in data-driven humanities, by using Big Data resources, information can flow in a paradoxical way. On the one hand, the provision of massive amounts of data can open paths for the exercise of autonomous individual and collective action. On the other hand, it has the power to exert undesirable influences on privacy and mass surveillance, threatening human autonomy by mechanically shaping messages, without ethical concerns, involving the reuse, repurposing, and recombination of personal data. Hence, DH might provide positive and negative fuel for the growth of a myriad of mechanical systems, with economic, epistemological, political, and ethical implications.

Final remarks

From the Complex Systems perspective advocated here, Digital Humanities should not be seen purely as an interdisciplinary field with challenges involving computers, humanists, and

modelling assumptions. Since, in DH, the digital escapes the walls of universities and enters our homes, it is necessary to incorporate ethical and epistemological concerns, considering the complex thinking perspective and the interrelationship between science and society. As pointed out, DH can help with the diffusion of reliable information and the use of computer modelling to analyse the digital shadow of humans, which might become more and more digital as the digital humanists import methods from the formal sciences. In a feedback loop, we might create our digital selves, which would be studied in the future by digital humanists using digital tools.

The complex thinking approach to DH allows the scientific investigation to undertake ethical statements from the outset: "The aim of complex system thinking is to produce the least possible uncertainty for responsible innovation." (Bourgine 2013). Considering that not everything that is technically possible is ethically defensible, a sound, genuinely interdisciplinary Digital Humanities should combine wisdom with the systemic integration of knowledge. This includes ethical, cultural, political, economic, and ecological contexts that should be treated as integrated and co-dependent.

In practical terms, a way of approaching interdisciplinarity in DH would be to develop an honest and humble dialogue among people of diverse backgrounds and expertise, with the possibility of developing research around problems and topics in common. This should include patiently sharing the weakest points of one's practices and seeking advice, together reflecting carefully on the possible consequences of the group's research for science and society. Encouragingly, some projects are emerging in the direction of collaborative efforts to think about the role of data and data science to society, examples being the movement 'Data for Black Lives'[7] and 'The Human Error Project'[8].

What could be the future of DH, and to what extent might its development threaten the degree of autonomy of human action? Considering the future of DH, the Complex Systems perspective alerts us to the difficulty of establishing precise predictions. However, insofar as autonomous action is concerned, it may (or may not) be threatened by the development of Digital Humanities, depending on the human capacity to understand the informational-technological dynamics of the social/ecological systems, and the conceivable ways of altering (when necessary) the processes that govern such dynamics.

Endnotes:

1. https://www.ibm.com/ibm/history/ibm100/us/en/icons/speechreco/team. Accessed in April 2021.
2. D'Ignazio & Klein (2020, p. 27) call attention to the qualitative *versus* quantitative dichotomy, stating that dichotomies usually hide underlying hierarchies: "[...] in this case it is that quantitative data can be incorrectly perceived as 'better' than qualitative data for being more objective, true, generalizable, larger, scale, and so on".
3. This was the case of a paper retracted from Nature in January 2021, which associated dishonesty with obesity: https://www.nature.com/articles/s41598-020-80528-2.pdf. Accessed in April 2021.
4. In other words, a sign can be characterized as a logical relationship that has three elements: (i) the sign-vehicle, which is determined by (ii) the object of a sign that determines (iii) an interpretant, i.e., the

possible effect of a sign which is, also, a sign (but a more developed sign). The sign-vehicle, object, and interpretant are signs that have different logical roles in the relationship that constitutes the sign.

5. For more examples of the mismatch between statistical tools, together with valuable insights into computational literary studies, see Da (2019).
6. In the case of GPT-3, it was trained with a massive dataset of 175 billion parameters, compared to the 1.5 billion parameters used for the second-generation model.
7. https://d4bl.org.
8. https://thehumanerrorproject.ch.

References

Alvarado, Rafael, and Paul Humphreys. "Big Data, Thick Mediation, and Representational Opacity". *New Literary History* 48.4 (2017): 729-749.

Bourgine, Paul. "Complex Systems Science: Lifelong Education by Research". *Notes from the conference on the XXV CIC – UNESP*, Brazil, 2013.

Breiman, Leo. "Statistical modeling: the two cultures". *Statistical Science* 16.3 (2001): 199-231.

Burdick, Anne et. al. *Digital_Humanities*. Cambridge, MA: MIT Press, 2013.

Callaway, Ewan et. al. "The push and the pull of digital humanities: topic modelling the 'what is digital humanities?' genre". *Digital Humanities Quarterly* 14.1 (2020). Available at <https://dhq-static.digitalhumanities.org/pdf/000450.pdf>. Accessed in 02/05/2021.

Collmann, Jeff. et. al. "Data management plans, institutional review boards, and the ethical management of big data about human subjects." *Ethical Reasoning in Big Data: An Explanatory Analysis*. Collmann & Matei (eds.), Cham: Springer, 2016, 141-184.

Da, Nan Z. "The Computational case against Computational Literary Studies". *Critical Inquiry* 45. 3 (2019): 1-34.

Drucker, Johanna. "Humanities approaches to graphical display". *Digital Humanities Quarterly* 5. 1 (2011): 1-21. Available at <http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>. Accessed 15/04/2021.

Floridi, Luciano and Massimo Chiaritti. "GPT-3: Its Nature, Scope, Limits, and Consequences". *Minds and Machines* 30 (2020): 681–694.

Gibson, James J. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton-Mifflin, 1979.

Gonzalez, Maria Eunice. "Autonomous action in complex mechanical systems: a real dilemma?". *Cognitive Science: Recent Advances and Recurring Problems*. Adams, Pessoa & Kogler (eds). Wilmington, DE: Vernon Press, 2017, 9-19.

Gray, Jim. "Jim Gray on eScience: A Transformed Scientific Method." *The fourth paradigm.* Tony Hey Stewart Tansley and Kristin M. Tolle (ed.). Redmond, WA: Microsoft Research, 2007, 17-31.

Grim, Patrick and Daniel Singer, "Computational Philosophy". In: *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.), 2020. Available at <https://plato.stanford.edu/archives/spr2020/entries/computational-philosophy/>. Accessed 05/04/2022.

Laney, Doug. "3D Data Management Controlling Data Volume Velocity and Variety". *META Group Research Note* 6 (2001).

Leonelli, Sabina. "What difference does quantity make? On the Epistemology of Big Data in Biology". *Big Data & Society* 1 (2014): 1-16.

Leonelli, Sabina. "What counts as scientific data? A relational framework". *Philosophy of Science* 82 (2015): 810-821.

McLuhan, Marshall. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill, 1964.

Macluhan, Marshall and Quentin Fiore. *The medium is the message: An inventory of effects*. Berkeley: Gingko Press, 2005.

Menant, Christophe. "Evolution of meaningful information generation through the evolution of life". *Gatherings in Biosemiotics 3*, Copenhagen 2003. Available at <http://philpapers.org/rec/MENEOM>. Accessed 08/03/2021

Menant, Christophe. "Computation on information, meaning and representations. Proposal for an evolutionary approach". I*nformation and Computation: Essays on Scientific and Philosophical Understanding of Foundations of Information and Computation*, G. Dodig-Crnkovic & M. Burgin (eds.), 2011, 255-286. Available at: <http://philpapers.org/rec/MENCOI>. Accessed 08/03/2021.

Misra, Ishan. et al. "Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", 2016, 2930-2939. Available at: https://arxiv.org/pdf/1512.06974v2.pdf. Accessed 03/04/2021.

Peirce, Charles S. (1958) "The collected papers of Charles Sanders Peirce". Electronic edition. Vols. I-VI, Hartshorne, C., Weiss, P. (eds.), 1931-1935. Vols. VII-VIII, Burks, A.W. (ed.). Charlottesville: Intelex Corporation. Cambridge: Harvard University Press. [Quoted as CP, followed by the volume and paragraph].

Pietsch, Wolfgang. "Big Data - New Science of Complexity", 2013. Available at <http://philsci-archive.pitt.edu/9944/1/pietsch-bigdata_complexity.pdf>, Accessed 03/01/2019.

Piotrowski, Michael. "Ain't no way around it: why we need to be clear about what we mean by 'Digital Humanities'". *Proceedings of the symposium "Wozu Digitale Geisteswissenschaften?"*. November 2019.

Roth, Camille. "Digital, digitized, and numerical humanities". *Digital Scholarship in the Humanities* 34. 3 (2019): 616-632.

Strickland, Elize. "OpenAI's GPT-3 Speaks! (Kindly Disregard Toxic Language)". *IEEE Spectrum,* 2021, https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/open-ais-powerful-text-generating-tool-is-ready-for-business Accessed 6/4/2021.

Uexküll, Jacob von. "The theory of meaning". *Semiotica* 42.1 [1940] (1982,): 25-82.

Uexküll, Jacob von. *A foray into the worlds of animals and humans: With a theory of meaning*. Trans. O'Neil, J. Minneapolis: University of Minnesota Press, [1934, 1940] 2010.

Weaver, Warren. "Science and complexity". *American Scientist* 36 (1948): 536-544.

Zoya, Leonardo Rodrigues and Julio Aguirre. "Teorias de la complejidad y ciencias sociales. Nuevas Estrategias Epistemologicas y Metodologicas". *Nomadas. Revista Critica de Ciencias Sociales y Juridicas* 30. 2 (2011) Available at: http://revistas.ucm.es/index.php/NOMA/article/view/36562. Accessed 04/04/2021.

Acknowledgments

This work was supported by FAPESP (no. 2020/03104-1) and CNPq (no. 309882-2019-2), Brazil. We would like to thank the members of our research group OpLaDyn (FAPESP no 2016/50256-0) and the members of GAEC (Academic Group of Cognitive Studies) for the inspiring discussions and feedback.