Pázmány Péter Catholic University

Faculty of Humanities and Social Sciences

Doctoral School of Literary Studies

Gábor Patkós

The Datafied Ontology and the Cultural Techniques of Digital Life

Doctoral (PhD) dissertation

Head of the Doctoral School: Prof. Dr. István Dobos Dsc

Supervisor: Dr. Zsolt Almási, associate professor

Budapest

2025

Pázmány Péter Katolikus Egyetem

Bölcsészet- és Társadalomtudományi Kar

Irodalomtudományi Doktori Iskola

Patkós Gábor

A digitális lét adatosított ontológiája és kultúrtechnikái

Doktori (PhD) disszertáció

A doktori iskola vezetője: Prof. Dr. Dobos István Dsc, egyetemi tanár

Az értekezés témavezetője: Dr. Almási Zsolt, egyetemi docens

Budapest

2025

Table of Contents

Acknowledgments

If someone were to tell me in the summer of 2021 how profoundly meaningful and significant my decision to pursue a doctorate at Pázmány would prove to be four years later, I probably wouldn't have believed them. But let's take it one step at a time:

When I first started school, my parents always told me that I had one job: to study. For a long time, this was not quite noticeable in my grades; it took me about 20 years to take it seriously. They probably did not think at the time that they would see this return in my dissertation at the age of 33. Luckily, when I needed it, they gave me all the help I needed to get this far.

Firstly, I would like to thank the PPCU Doctoral School of Literary Studies for giving me the opportunity to pursue my doctorate here. I am grateful to my PhD supervisor, Zsolt Almási, for his sustained confidence in me since our first conversations that this project could be successful. Speaking of supervisors, this is the place to sincerely thank Miklós Péti, my BA supervisor — who became a trusted mentor and friend since then — for his continuous support and guidance throughout the past 15 years.

During my doctoral studies, I have grown enormously both professionally and intellectually, and the years I have spent at PPCU have provided me with everything I needed to build a foundation for my professional career. I would like to give special acknowledgement to the courses taught by István Dobos, Kornélia Horváth, Kinga Földváry, and Veronika Schandl, as without what I learned in those courses, this dissertation would certainly not have been possible: much of the theoretical background and related ideas were born during those sessions.

I am grateful to the doctoral school not only for finding a professional home, but also for finding a family in the most literal sense of the word: I met my wife, Petra, on the first day of the programme and later our two wonderful children were born at the crack of dawn just before a school day. As my partner in life and in thought, if there is anyone without whom this dissertation could never have been possible, it is you.

Finally, I would like to dedicate my dissertation to my grandfather, who was never involved in academia, but whose gentleness, humility, peacefulness with those around him, and very importantly a deep understanding and the ability to clearly explain and demonstrate how things work set an example that shaped my entire outlook on life.

1. Introduction

"*Against that positivism which stops before phenomena,*

saying "there are only facts," I should say: no,

it is precisely facts that do not exist, only interpretations"

- Friedrich Nietzsche, Notes 1888

Today, the discourse of attempts at defining our 21st century digitalized society is almost commonplace. Several terms have been coined in recent decades to describe this phenomenon: post-industrial society (Touraine), postmodern society (Lyotard), postcapitalist society (Webster), knowledge society (Toffler), digital capitalism (Schiller), information capitalism (Fitzpatrick), fourth industrial revolution (Schwab), AI revolution (Urban), the list goes on. However, it seems to be an empirically verifiable fact that more and more of our lives are moving into the digital space dominated by information networks. The question naturally arises as to what the real essence of this change is, what are its starting points and underlying assumptions, and what phenomena are merely symptoms of change.

A prominent sociologist and theorist, Manuel Castells, worked extensively on the concept of information society in his highly ambitious, 900-plus-page trilogy *The Information Age: Economy, Society, and Culture*. In the first volume, originally published in 1996 titled *The Rise of the Network Society*, Castells articulates the logic of the network society, by which he means that:

> As a historical trend, dominant functions and processes in the Information Age are increasingly organized around networks. Networks constitute the new social morphology of our societies, and the diffusion of networking logic substantially modifies the operation and outcomes in processes of production, experience, power, and culture. While the networking form of social organization has existed in other times and spaces, the new information technology paradigm provides the material basis for its pervasive expansion throughout the entire social structure. (500)

Castells thus identifies the totalitarian spread of networks as a consequence of the information technological paradigm. He seeks to illustrate the effects of this paradigm by distinguishing between the terms "information society" and "informational society". He argues that most societies throughout history were information societies as far as information and communication played a significant role within them. However, the term "informational society" is defined as "a specific form of social organization in which information generation, processing, and transmission become the fundamental sources of productivity and power because of new technological conditions emerging in this historical period." (21).

Jan A.G.M Van Dijk, sociologist and researcher of communication and media studies, developed another important concept in 1999 what he calls the network society as the social system of the 21st century, which he defines "as a social formation with an infrastructure of social and media networks enabling its prime mode of organization at all levels (individual, group/organizational and societal)" (*The Network Society* 20). However, it is also important to note that these networks are not only present, but also fundamentally determine the functioning and organization of society. In Van Dijk's words, "these networks increasingly link all the units or parts of this formation (individuals, groups and organizations). In Western societies, the individual, linked by networks, becomes the basic unit of the network society" (*The Network Society* 20). In a later paper, however, the author tries to clarify and nuance Castells' approach by insisting that "Society still consists of individuals, groups/pairs and organizations. Of course, they form external and internal relations, but these relations do not equal society ("One Dimensional" 133).

Van Dijk and Castells both present a convincing and thorough account of the social rearrangements caused by digitalization, but their horizons remain within the boundaries of the episteme. Put simply, they do not examine digitization itself or the ways in which digitization comes about, but rather its visible and describable consequences. In my opinion, it is precisely for this reason that the epistemological approach is the most appropriate for this enterprise, which aims, following Foucault, to sketch the outlines of the archaeology of digitality.

In his early work titled *The Archaeology of Knowledge*, Michel Foucault introduces his concept of the episteme, the proceeds to examine different historical periods from the perspective of epistemology. In his conceptualization, episteme is a fundamental epistemological field that is a condition for the emergence of knowledge. "This middle region, then, in so far as it makes manifest the modes of being of order, can be posited as the most fundamental of all" in the author's words, "it is here that it appears, according to the culture

and the age in question, continuous and graduated or discontinuous and piecemeal, linked to space or constituted anew at each instant by the driving force of time, related to a series of variables or defined by separate systems of coherences, composed of resemblances which are either successive or corresponding, preface xxii organized around increasing differences, etc." Foucault describes his own program for the archaeology of the human sciences as follows: "it is rather an inquiry whose aim is to rediscover on what basis knowledge and theory became possible; within what space of order knowledge was constituted; on the basis of what historical a priori, and in the element of what positivity, ideas could appear, sciences be established, experience be reflected in philosophies, rationalities be formed, only, perhaps, to dissolve and vanish soon afterwards" (Preface, XXIII). It is important to note, then, that we are talking about dynamic discourse networks that are transitory, constantly evolving and changing. The starting point of my dissertation is this Foucauldian assumption, and my thesis is that the phenomenon of digitalization and datafication can be understood as the ruling episteme of today's time. Consequently, all further developments, such as social media, video games or even AI-based technologies are the consequence of entering this episteme. My aim, therefore, is to develop and explore theory of the episteme of the digital, which I call digitality throughout my dissertation.

In some ways, it may seem that the humanities have little to do with theories of digital technologies. Although it is increasingly integrating its tools and methodologies, especially in digital humanities, the concept of digitality itself is less prominent as an object of study in the humanities. Robert Hassan, in his work, *The Condition of Digitality*, referring to David Harvey's seminal work, sets himself precisely this task, introducing the concept of digitality outlined as a program intended to

> prioritize instead a humanist understanding of the processes of a machine, a logic, that has not only rapidly colonized every part of the inhabited planet, but has also suffused the consciousness of almost every person within it in terms of his or her engagement with each other through networks of communication, production and consumption: I call it digitality. (1-2).

Hassan thus makes it clear that by digitality he does not simply mean an economic or social structure. Instead, he considers digitality as a state of existence with far-reaching cultural, even epistemic consequences. Adapting Hassan's perspective, I am convinced that we must extend the paradigm of the humanities to digitality, because it is the science that can study digitality itself, digital objects, in the right depth and from the right perspective.

At this point, I believe it is worthwhile moving to discussing the subjects, objectives and methodologies of the humanities. Matthias Freise, following Dilthey in his search for an interdisciplinary "common ground" in the humanities, finds it in the study of relations (xviii). Freise claims that the point of the humanities is to interpret relations between different objects (vii). However, when we start to treat these relations as objects in their own right, we lose sight of the goal (xx). Freise thus identifies dialogicity as the most fundamental methodological approach to humanities, in the broadest sense, rather than hermeneutics or semiotics in the classical sense. As the author puts it, "The acknowledgement of relationship is impossible without interpretation. Interpretation is the readiness to relate. If I abstain from interpretation, I refuse relationship. (xviii)."

Today, these links are fading, we are in a kind of neo-positivism, the main trend of which is the adoption of a data-driven, "big data" approach. This perspective threatens to define everything as an autonomous, "given" object. As Chris Anderson, editor of the Wired Magazine wrote in a 2008 editorial, titled *The End of Theory*:

> This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves. (no pag.)

It is doubtful, that the numbers speak for themselves. Anderson's words on the other hand, clearly do: in the 21st century there is no need for interpretation, no need for understanding. All we need is measurement, calculation and analytics. This is also clearly evident in the thinking of Lev Manovich, who in his program calls on the humanities to make the data themselves, the data sets, the objects of cultural analysis, rather than the study of relations (70). Interpretation is relegated to the background in relation to the "numbers" that can be known and taken as objective, and the freedom of the human subject is limited to the choice between the alternatives offered, to consumption.

The general aim of my dissertation is to redefine and reposition the possibilities of human interpretation in an information- and data-centered digital life. I intend to achieve this through proposing a theoretical framework, the episteme of digitality, through which we are able to contextualize and interpret intermedial cultural and epistemological phenomena that are already inherent in digitality. These phenomena, which can be considered as cultural techniques

of digitality, will be analyzed in the form of case studies according to the criteria of the framework developed.

2. The Conceptual Framework of Digitality

The digital revolution is a profound change in science and society brought about by pervasive digital technology. It involves a sweeping change in the way knowledge is created, disseminated and understood. The German Research Foundation (DFG) defines the digital turn broadly as encompassing "all relevant changes and effects of an epistemic, ethical, legal, technical, infrastructural, organizational, financial and social nature resulting from the development and use of digital technologies in the sciences and humanities" (DFG, German Research Foundation). In the humanities, this means the shift from purely analogue practices, such as print-based philology and close reading to quantitative methods, digital tools and even big data approaches. Boyd and Crawford write of the "Big Data" research perspective that "reframe[s] key questions about the constitution of knowledge, the processes of research, how we should engage with information, and the nature and the categorization of reality" (3). This transformation has sparked debates about continuity and change: some observers herald a shift in research paradigms and even the "scientification" of the humanities, while others note that most humanities scholars now work with a hybrid approach, combining traditional and digital methods. (C2DH, no page number). In fact, a whole series of lectures have been devoted to discussing these issues between 2019 and 2021 at the Centre for Contemporary and Digital History at the University of Luxembourg.

It is important to note, therefore, that the digital turn has dramatically transformed the way knowledge is produced in the humanities. In many cases, digital tools and data-driven methods of other fields have transformed research practices and epistemologies of humanities as well. At the center of these transformations are processes within literary studies, an often-discussed example is Franco Moretti's concept of distant reading, which is the analysis of large literary corpora using computational methods. Such approaches, a leading aspect of the digital turn, have "influenced literary studies" by shifting the focus to aggregate samples and triggering debates on genre, authorship and style on a new scale (Primorac et al, 16). More generally, across disciplines, scholars are now mining text archives, mapping historical data and visualizing cultural trends, asking new kinds of research questions that only digital methods can answer.

Furthermore, the digital turn has led to crucial reflections on epistemology in general, prompting us to explore the nature and validity of knowledge, especially in terms of different epochs of technology. Traditional humanistic scholarship emphasized depth, context and

qualitative interpretation, often presented in linear narratives, such as a monograph, often considered the most important requirement of a habilitation. On the other hand, the influx of digital methods introduced quantitative scale, concerns about reproducibility, and new ways of reasoning, such as interactive visualization or database-driven analysis. Scholars of humanities and social scientists needed to find answers to questions and concerns to new questions: How do algorithms and data shape what we consider evidence? How can we ensure that interpretation keeps pace with computation?

In response, I find it important to stress that digital methods need to be integrated critically, not just for novelty. For example, the early enthusiasm that computers would bring neutrality and objectivity to humanities research has been tempered by an awareness of bias and context. Lorella Viola notes that the digital turn initially "assumed to be, and often advertised as being neutral, fair and accurate" creating the illusion that technology could be used without changing established models of individual disciplines, but in reality, often exacerbated complexity and inequalities, showing how illusory these promises were (15). This critique emphasizes that digital tools are not value-free and that knowledge production in the humanities still requires humanistic judgements about meaning, not to mention ethical and social implications.

Naturally, the ways of disseminating and communicating science have also changed. In the digital humanities, scholars often share knowledge not only in print, but also through online journals, often following a continuous publication model, or open access repositories. Podcasts, videos or other multimedia formats are also gaining traction. Forms of research products are also undergoing change: it is possible to produce digital databases, websites or even software as scholarly outputs. These new formats broaden the audience and usability of humanities research but also require new standards of peer review and preservation. Most of these developments are becoming standard due to the general prevalence of digital infrastructures.

The digital turn has also fostered an ethic of openness: open data, open-source tools and public scholarly activity are becoming more common. As information is transparent, scholarly work done in digitality will have the tendency to be more accessible. The boundaries between academia and the public can become blurred when archives or analyses are made available online, following the more general cultural shift in knowledge exchange and digital communication in the information age. Although there are tensions between quantitative approaches and interpretation, scholars tend to argue for a more inclusive paradigm where digital techniques complement, rather than replace human research.

This is an especially important consideration as the hermeneutic approach itself is beginning to re-emerge. In a 2024 article, Couldry articulates a call for a hermeneutic of today's anti-hermeneutic landscape of algorithmic influence. The article highlights Martín-Barbero's work on cultural complexities, placing the critique of a monolithic depiction of culture at the center of attention. Couldry then moves to describe the challenges of the "platformed transformation" of popular culture: commercial companies can now not only produce cultural content, but the very terrain on which cultural life is lived (6). The article posits the question, "How do we make sense hermeneutically of social spaces which have been created precisely so that what goes on in them can be tracked, nudged and, in important respects, managed by the platforms that built those spaces?" (7) which is, in some ways the exact same question I ask regarding humanities. Popular culture in digitality as described by Couldry, also provides a point of connection for my project of theorizing digital life.

Following the idea of anti-hermeneutic culture, it is important to consider the various posthumanist positions that are increasingly common in theories of digital technologies in the 21st century. Although theories of posthumanism are very different, a common feature shared by most perspectives is to question and revise the traditional human-centered, so-called Anthropocene worldview in the light of information technologies, artificial intelligence and biotechnology. According to posthumanist thinkers, the digital turn is accelerating the transcendence of the exceptional status of the human being and decentralizing the human subject in favor of networks of humans and non-humans. For example, Rosi Braidotti, a proponent of critical posthumanism, argues that contemporary "posthuman knowledge" goes beyond the old ideal that humans are the measure of all things (1-2). Her article posits a philosophical foundation of "intelligent and self-organizing matter", which implies that cognition and agency are distributed across sets of people, machines and things (4). In this framework, the cognizing subject becomes "relational, embodied and embedded" rather than a disembodied rational mind (12).

Along the same lines, Donna Haraway's cyborg theory or Bruno Latour's actor-network theory provide important insight to consider how digital tools, algorithms and data processes intertwine and interact with humans in the production of knowledge. Similarly, posthumanist positions encourage the humanities to question traditionally accepted notions of authorship, consciousness and authority when, for example, algorithmic systems are involved in analysis or when archives are "thinking" through metadata. Frameworks of posthumanism thus intend to reshape the epistemology of the humanities, extending it to concepts such as machine agency,

network dynamics, and in general argue for the dissolution of clear boundaries between human culture and digital technology. Although this is an important consideration for philosophy and may shift researchers' attention toward technology, it is a slightly different direction from the goal of this dissertation. The primary aim of my project is to examine human agency of interpretation and reception, and the impact of data-driven ontology on it. Therefore, my question ultimately is always how an algorithm affects the human subject, not if an algorithm is capable of acting.

Visual and computer tools, as well as theories of posthumanism that focus on non-human cognition, do not provide productive continuity for human interpretation in the face of media changes. In many ways, Ludwig Pfeiffer's medial anthropology proves to be a useful approach to achieve this, preserving the possibility of focusing on the human being rather than posthuman theorizing. However, Pfeiffer states that "I do not yet consider it likely that electronic (multi)media and the entertainment possibilities of cyberspace will have a decisive impact on the human nervous system" (26). If that was not the case in 1990s, which is somewhat doubtful, neuroscientific research of the 2010s clearly shows the opposite, which Picchione illustrates in his study with numerous examples, including the negative impact of digitality on social relationships, the links between attention deficit and digitality, and poor performance in reading comprehension on screen (19-20). I would argue that this trend seems to indicate that we are indeed entering a network of discourse on digitality, which may still demonstrate the principle of equivalence, but the dialectics of the changing episteme are becoming increasingly apparent.

2.1 Discourse Networks

Western philosophy has been thinking about the logos for over two thousand years and has a highly influential tradition of Aristotelian notions of the "rational" or "language-possessing" animal. Human communication, in this conception, is not only a response to environmental influences, but also an awareness of what is right and wrong (Pol. 1.1253a). However, this line of thought is increasingly problematized by the many achievements in the field of biosemiotics (Brentari 331). It seems a plausible axiom, however, that communication is a basic anthropological necessity, but perhaps it is not too much to venture that it is a fundamental attribute of all life. If we accept this, we arrive at the conclusion that language can ultimately be most accurately described as a (bio)technology of communication evolved from one of our anthropological needs. It follows that our communicative technologies are responses to our anthropological needs, which differ in the nature of their mediality. Meyrowitz's theory of the medium reflects precisely this, in that different media, because of their different characteristics, privilege certain interactions and retard others (50).

At this point, it is useful to discuss the foundations of Kittler's media theory, and the author also has an important starting point in the archaeological orientation of early Foucault, especially in *The Order of Things*, in which he identifies different epistemological epochs, which he then analyses according to their internal workings. Kittler identifies these epistemes as discourse networks, which he defines as "technological and institutional networks that enable cultures to organize, store, and process data" (*Discourse Networks* 368). Kittler thus sees the change in the discourse networks built upon them as the driving force of epistemic change. The primary cause of the change in discourse networks, according to Kittler, is war. We can therefore say that the evolution of the media is an expansion of war and hostility, with the media following the escalation model of war (*Valhalla* 6).

However, it is not only Foucault's work that Kittler's media theory is inspired by. Yuri Lotman's model of cultural semiotics also works with the terms explosion — gradual development. Gradual development is a period in culture when there is a basically predictable, modellable rationality, followed at some point by an unpredictable explosion. However, these periods not only alternate, but can also exist in parallel: different layers of culture can exist at different speeds, in different states (Lotman 62). These processes perform important functions that are characteristic of any complex semiotic system; some processes promote innovation,

others continuity, their effects describe as antagonistic, indeed, according to Lotman, “the struggle between them can be described by the category of total war” (62).

Similar to the authors mentioned above, but perhaps in a slightly more humanistic way, the metaphor of war appears in Chantal Mouffe’s political philosophy. Mouffe argues that in a truly pluralistic system, the “other” becomes not an enemy but an adversary. There is an “antagonistic” relationship between enemies, but only an “agonistic” relationship between opponents. In this relation, each individual seeks to assert his own interests, but it is also recognized that the opponent has an equal right and opportunity to assert his interests (Mouffe 756) In fact, it is a question of mutual acceptance of opposition, of a kind of consensus based on conflict (ibid). If one accepts Kittler’s argument that war is the a priori of the media, the assertion of the agonistic principle may bring us one step closer to a truly interdisciplinary, more democratic episteme.

I would also like to use a network approach to describe the technological, medial discourse system of digitality. Marshall McLuhan, in his *Understanding Media*, argues that the content of any medium is always just another medium (10). This seems to be fundamentally true, but it needs to be complemented. Rather than a linear genealogy or a chain of development, my thesis, like Kittler’s (and partly Pfeiffer’s), is that the various media can be described primarily in terms of complex systems theory. More specifically, a scale-free network based on the Barabási-Albert model. The vertices of the network I am proposing are the media, and the edges between them express the relationship between which other media are “contained” by a given vertex. In simple terms, the edges between vertices represent the intermediate set of relations.

The uniqueness of the Barabási-Albert model is that it creates a continuously growing network, or so-called preferential attachment. This means that nodes with more edges connected are more likely to have new edges. Most vertices are expected to have only a much lower number of edges (Albert and Barabási 27). It is important to note that one of the earliest starting points for the research of the author duo, Albert-László Barabási and Réka Albert, is the study of the World Wide Web. It is no coincidence, therefore, that the model seems highly suitable for studying the medial medium of digitality. As the authors write about research on the Internet (among other things), I would include my own endeavor among these efforts: “special attention has been given to a few networks of real, outstanding technological or intellectual importance. In these studies, the aim is to develop models that go beyond basic growth mechanisms and

incorporate the specific and often unique details of the system in question (Albert and Barabási 50)".

This approach, while bearing McLuhan's most important theoretical insights, avoids the criticisms of technological determinism that are often mentioned and have now lost its practical relevance due to the different socio-economical and technological environment at present. Instead by describing a dynamic system we can identify tendencies that are capable of changing with the medial environment around is. Furthermore, it is able to integrate a number of additional approaches to further justify the theory:

Firstly, as network theory is inherently interested in relations within complex actors in a system, this approach appears to be methodologically more suitable to be analyzed from a semiotic perspective. As opposed to hierarchical or more rigid formalist categories, this allows important terms related to Lotman's cultural semiotics, such as semiosphere or autocommunication, to be imported into digital media theory.

Second, as Pfeiffer puts it, we can identify media configurations in the complex system in a variety of ways: "The media, at least in tendency, appear at least in the form of implicit or implicit combinations (in the mode of intermediality or hybridization in the meaning of the world according to McLuhan)" (22). The inherent intermediality (and possibly interdisciplinarity) of culture that Pfeiffer emphasizes lends itself to complex systems. The relationality of knowledge and culture are important elements of poststructuralist theory as well, as exemplified by Deleuze and Guattari's concept of the assemblage.

Third, we could be able to create and even visually represent a space in which the possibility of examining the relations of media is opened up, and even to give space to human inclinations to action, in the words of Tamás Pólya: "when social change occurs through the emergence of a new medium, the ultimate cause of change is human inclination, not technology: the technical medium only opens or closes off dimensions of action, it enlivens or cools down already existing human intuitions" (70).

However, the simple placement of language or logos, with all the complexities and problems of the term, in this system still seems to be a frivolously reductionist procedure. However, the typology I propose may still be able to clarify the role of language. If one simply assumes that language is a central node to which almost all other media are connected, the privileged place of language as a preferential node immediately begins to emerge. The role of a given language — in Saussurean terms langue — can be clarified with far fewer reservations.

The same applies to spoken language — or, also from Saussure, parole. I would trace the questions concerning the role of language in my theoretical framework back to three reasons:

First, language is fundamentally embodied, it comes into being directly in the embodiment. Language is the body itself, with which we are in direct contact without any special mediation. In other words, language itself is the mediation, a kind of pure mediality ("Afformative, strike" 1155). The other technical media are either connected to our senses or to language. Writing does not directly use our thoughts as a medium, but only the thoughts we produce through language. I do not mean to refer to the logocentrism of spoken language or the secondary nature of writing, but to the inherent condition of language as a technology, whether writing or otherwise. In the same way, the regular system of visuality or the frequency of sounds obviously does not presuppose spoken language but nevertheless functions as a kind of language. The mechanisms of autocommunication or the semiosphere are likewise tied to a kind of language.

Second, language appears to be both pure materiality and immateriality. All other media can only be understood in relation to these, on this scale, but language is a dichotomy. By having direct access to language as the central element of our thinking and experiencing it in embodiment, it would be more difficult to present a more material medium. Yet, at the same time, the multiplicity of theories of language and our difficulty in grasping the logos attest to the fact that "language itself" is even more difficult to grasp than even the essence of information. Hamacher's afformative theory of language points out that language is really present precisely in its own absence, in its suspension by itself (Balogh 105), which in turn testifies to an insuperable immateriality.

Third, the invisibility of the thinking process, leading to the distinction between language and thoughts. Although language is intrinsically involved in the process of thinking, perhaps even more directly than other media, we cannot know anything about it outside of language, since our invisible and inexpressible thoughts can only become visible through language, in language, and through language. Thus, all communicative acts ultimately take place in (some kind of) language.

Another big challenge is digitalization, or digital media. From a media point of view, I think it is possible to define digital as a separate medium and then place it in a complex system of media. In this case, digitality also appears as a preferential node in the network. However, a second way is also possible. If we describe media technologies as a multi-layered network,

digitality appears as a separate layer in which analogue media appear in a specific way, but in a way that parallels their analogue counterparts. I do not see this question as settled at this stage of the research.

In my dissertation therefore, the only thing that I believe could be compared to the rise of digitality and the epistemological consequences of data is the emergence of literacy. In my dissertation I will therefore refer to the rise of literacy on several occasions, drawing parallels with the spread of print culture. It is important to note that, while writing as a technology has been used for thousands of years, it was not until the second half of the 20th century, with the advent of digitality, that the theory of writing as a mediality began to be developed.

This is, of course, no coincidence: N. Katherine Hayles argues in relation to the poststructuralist shift in critical theory that it came about because the absence/presence dialectic characteristic of structuralism had already lost "its grounding power for discourse" and thus became visible (82). According to Hayles, then, it is not that critique brought about the epistemic shift, but rather that the critical shift could only have come about because a different dialectic had already been set in motion. If we follow this logic in relation to the linguistic turn or even to grammatology — the science of writing, which Derrida notes in the 1960s as "showing signs of liberation thanks to considerable efforts" (12-13) — we can assume that these shifts are also the consequence of the entry into the information age and are in fact about a dialectic that has been superseded. The question then arises whether, if I am talking about the theory of digital media, we have moved beyond it. It is possible that with the rise of artificial intelligence we are already in a new paradigm, but the epistemic shift has not yet been revealed.

At the same time, it can be seen that literature in the narrow sense (the science of written texts) has lost its privileged role as the hegemony of writing as a medium has been broken, and literature is increasingly present alongside other media (see Földváry 2022). Unfortunately, literature (and the humanities in general) often characterize this process with a pessimistic, sometimes even defeatist discourse, but it is important to recognize that these processes are not negative in themselves, but are merely natural consequences of changes in discourse systems. Like all change, they are also opportunities. Literature and cultural studies are at an important moment when digital re-ontologization can enable them to move forward and renew themselves.

I would like to make it clear, however, that this is by no means a self-abandonment: the methods and traditions of literary studies, philology and the humanities more generally are the best suited

(among others) to interpreting cultural phenomena in the digital world. The renewal that I am calling for lies in the broadened perspective of the focus and objects of study through the theories of literary studies. Nevertheless, it is important to recognize that, at the moment, it may well seem that the human-centered possibilities in the infosphere are becoming increasingly narrowed. The same is true of the role of language. The human, the linguistic, may seem to be moving out of the center, from its prominent role to just one of many. Seen from this perspective, the state of digitality may indeed be more accurately described by the posthuman perspective, but the language of the posthuman is not human language, but code ("Condition" 90).

We have seen that just as language became problematic, information becomes problematic by the end of the 20th century. I propose to conclude that we are witnessing a critical transformation of knowledge production. However, in order to continue my analysis, I need to turn to discussing the relationship between culture, knowledge production and technology.

2.2 Technology and Culture: Determinism and Constructivism

At this point, I think it is important to expand on the relationship between technology and culture. A fundamental fault line in thinking about the relationship between technology and culture is between the theories of technological determinism and social constructivism. Technological determinism is basically the idea that the development of technological means is in itself a determining force that shapes and molds human society and thought (Wyatt 166). The roots of this view can often be traced back to the work of Marshall McLuhan and the Toronto School. McLuhan's famous thesis is that "the medium is the message", that is, the medium of communication itself shapes social message and experience. McLuhan divided human history into eras based on the dominant media of communication (tribal oral, literate, print and electronic), indicating that in each era the dominant media technology determines the way we perceive and think (Zhu 85).

Marshall McLuhan and his disciples also argued that technological innovations, such as book printing, created new structures of perception and social arrangements. McLuhan argued, for example, that the Gutenberg Galaxy, the era of the printed book, promoted linear, analytical thinking and contributed to the emergence of nation states and the industrial revolution (Gutenberg Galaxy). Thus, for the proponents of technological determinism, printing was not merely a new way of recording knowledge, but a transformer of the entire cultural ecosystem. Pólya (2011) writes in detail about the problems with this theory.

Yet, one of the most prominent exponents of technological determinism is Friedrich Kittler, who does not even attempt to place the possibilities of human action in his theory. Instead, it is precisely the "expulsion of the spirit from the sciences of mind" that he has set as his banner, the title of a volume he edited in which his early lectures are included ("Austreibung des Geistes aus den Geisteswissenschaften"). His radical thinking can best be summed up in the opening phrase of his volume Gramophone, Film, Typewriter: "The media determine our situation." (Gramophone, XXXIX), suggesting that the entire medium of human experience and understanding is given form by the media technologies available to us — just as the archaeology of knowledge in the Foucauldian sense shows that the thinking of every historical age is determined by the technical conditions of seeing and speaking. In Kittler's view, then, the infrastructure of technical means is a quasi-transcendental condition for the experience and

knowledge of a given age. In this conception, media and technology are not neutral mediators, but active agents that fundamentally determine what and how one can know in a given age.

Critics of the deterministic view, however, stress the opposite. Constructivism is more of a catch-all term, but the perspective of the various schools of thought is that the development of technology and its social effects are ultimately a function of human choices and social forces. Their basic premise is that it is not technology that determines society's ways of producing knowledge and culture, but rather society that determines which technologies it favors, how it develops them and what it uses them for, making the question of human responsibility and control crucial. One of the most important proponents of this perspective, Andrew Feenberg's critical theory of technology, basically agrees that our technological tools are not neutral: they favor certain ends and inhibit others. Feenberg emphasizes that the direction of technological development can be shaped and democratized: society has a say in the values and interests that designers incorporate into new tools (Feenberg 3). In Feenberg's words, "Technological design is therefore an ontological decision with political consequences" (3). Policy makers and the communities that oversee them have a responsibility to choose among possible directions for technology that serve human well-being and justice. Feenberg thus rejects both the naïve optimism that technology will bring progress on its own and the deterministic attitude that technology is an uncontrollable process: technology, he argues, is a socially shaped space in which democratic participation is key (3).

As early as the 1980s, Langdon Winner pointed out that we tend to take an uncritical, "sleepwalking" attitude to new technologies, the consequences of technology only becoming clear long after the technology has been used. In Winner's words, "In the technical sphere, we repeatedly enter into social contracts whose terms only become clear after they have been made" (53). Winner also warns that, although we have since the Enlightenment come to regard technology as a fundamental engine of progress, this belief often prevents us from anticipating the real social impact of technological innovations. Only in hindsight do we discover at what cost and with what side-effects a given technology has transformed our lives (50).

However, Winner also stresses that although we use technology as a tool at first, once it becomes part of our everyday lives, we see it more as a tool. "As technological tools, methods and systems become woven into the fabric of everyday life, they shed their mere tool-like nature and become part of our humanity," Winner says. In other words, technology is so integrated into human activity that it is inseparable from it: just think of how the smartphone or the internet

has become almost an extension of our bodies and minds in the 21st century. In this way, Winner nuances the deterministic picture: even if technology primarily creates new possibilities, we ourselves become the "essence" of technical systems, as our way of life adapts to them.

The debate between technological determinism and social constructivism sheds light on whether, in the relationship between technology and knowledge, it is the tools that determine human thinking or human factors that determine the ways in which the tools are used. It is likely that the reality lies between these two extremes, but it is not nearly so simple. As Winner points out, by the 20th century at the latest, it had become abundantly clear — especially in the wake of the devastation of the two world wars — that technology is not merely a servile tool but an active force that transforms culture and even human existence itself. (Source) At the same time, the impact of technology is never automatic: it is always embedded in social relations, and it is these relations that decide whether an invention becomes an atomic bomb or a nuclear power plant, whether a printing press distributes fiction or propaganda pamphlets. To understand this dialectical relationship, however, we need to go beyond the dichotomy of mere instrument and man — or, more precisely, user — and consider a third possibility.

To resolve the above debate, it is worth considering an approach that treats technology and man not as opposing factors, but as inherently inseparable, intertwined beings. This idea emerges from Bernard Stiegler's early philosophy of technology, which radically rethinks the Aristotelian episteme/techné distinction. Stiegler argues that man is fundamentally dependent on artificial extensions, or prostheses, so that technique is not merely instrumental but fundamentally determines human existence (16.) He thus questions the classical opposition in Greek philosophy between techné (practical knowledge, craft) and episteme (theoretical knowledge), which, according to Stiegler, are closely linked and mutually condition each other in human development, since both arise from man's inherent imperfection (1.)

Stiegler develops the thesis of man's "originary prostheticity" through the myth of Prometheus and Epimetheus, found in Plato's dialogue Protagoras, which Stiegler quotes at length (187-188). According to the myth, it was Epimetheus' task to distribute faculties among the creatures, but by the time it was man's turn, he had mistakenly distributed all the faculties, leaving man without natural faculties, and imperfect — a mistake that embodies man's inherent lack and vulnerability (188). When Prometheus was confronted with Epimetheus' mistake, there was nothing he could do. To help the fallen humanity, he stole from the gods Hephaestus

and Athena the technical knowledge and fire needed to use that knowledge so that man could survive. Prometheus thus provided mankind with technicity, an artificial, prosthetic way of being that fundamentally defines man's nature and temporality, characterized by foresight, anxiety, and reliance on external memory and tools (193). Thus, the myth symbolizes the dual structure of human existence: the original omission of Epimetheus (hindsight) and the compensatory foresight of Prometheus, which Stiegler presents as fundamental elements of the conception of humanity as an inherently technical or prosthetic being (193-194).

Mankind is therefore not biologically specialized in a particular environment; his survival and development have always been made possible by the use of technical means — fire, tools, language, writing. Thus, technology is not an external addition to man, but a fundamental, inescapable part of the conditions of human existence. The same idea is expressed by Jacques Derrida in an interview: "there is no natural, original human body: technology has not been added to man from outside, as an afterthought, like some alien body" ("Rhetoric of drugs" 244).

In light of this, Stiegler argues that there is no sharp boundary between man and technology: the human body and intellect have also evolved through coexistence with technical means. "The technical is not a mere extension of the human body, but rather a constitution of the human body. For man, then, the technic is not merely a means but an end" (152-153). In the same way, the invention of writing was not merely a technological innovation, but fundamentally changed human thought, memory and culture — so much so that without literacy we would have no idea today what we would consider knowledge. The same can be said of book printing or digital computers.

Therefore, for Stiegler, it is pointless to ask the question "whether technology defines us or we define technology". Man and technology always exist in a state of mutual genesis: man creates technology, and technology reflects back on man and shapes what we call knowledge. The successive historical ages — for example, the age of the manuscript codex, the age of the printed book, and then the digital age — each represent a particular system of knowledge and culture that has emerged from the co-evolution of the human spirit and the technical means.

Stiegler's approach thus goes beyond the determinist/constructivist debate: he neither sees technology as an omnipotent, external force, nor idealizes man as an absolute autonomous

subject. Instead, it describes the relationship between knowledge and technology as part of a dynamic feedback chain. In this chain, every new technology (e.g. a printing press or an artificial intelligence algorithm) is a product of human needs and imagination, but once it is created and diffused, it creates new conditions for human thought and culture. This is how the "knowledge economism" of different ages have evolved: for example, the humanist knowledge mediated by book printing in the Renaissance, or the information knowledge mediated by digital networks today.

2.3 The Cultural Impact of the Digital

In the information society, we increasingly use a whole repertoire of terms related to understanding. Terms such as critical thinking, filter bubble, fake news, fact-checking have become buzzwords that we talk about a lot, yet they have no productive content. I see three basic reasons for this. First and perhaps most importantly, there is a complete misunderstanding of how digitality works: what is imagined to be the faults, problems or symptoms of digitality is in fact the inherent nature of digitality as a discourse network. From this point of view, it is no wonder that we cannot find solutions to these problems.

The second problem I see is that in digitality, the task of understanding is imposed on the individual without providing the necessary tools and opportunities. If one accepts Han's argument, one could say that communication and information in digitality has multiple "purposes" and functions, but that facilitating real understanding is not one of them. Sybille Krämer writes of interactions in digitality: "But behind the networked screens, the user is hidden from the rhizome-like, multifarious landscape of interacting protocols, machines and data. This deep region, destroyed by the cultural technique of flattening, returns as an invisible, uncontrollable space, inaccessible to the user (93).

At this point, the problem is linked to point three, which is that we increasingly want to leave the act of understanding to the digital systems themselves, and even expect it from the systems themselves. I am not just referring to artificial intelligence, but to any basic information system function that is a fundamental element of digital interfaces. Examples include search and filtering functions; ranking and indexing algorithms; or even just editing software. It is clear that these digital crutches, through mechanisms hidden from the user, facilitate interaction but not the act of understanding. The problem is therefore twofold: we want to facilitate, speed up or, in some cases, even get away with the act of understanding — and when this fails, we blame the system or the user instead of the problematic dimensions of the interaction.

Gadamer identifies the medium of hermeneutics in language, stating in his hugely influential Truth and Method that "language is the medium in which mutual understanding and agreement on the matter between partners takes place" (269).It is interesting to parallel this idea with Lev Manovich's cultural analysis, whose central starting point is the argument that human language cannot describe digital culture in its entirety. Instead, we must turn to "big data," data, algorithms and visualization, which he seeks to articulate as a kind of language in its own right,

inherently better suited to such an enterprise (*Cultural Analytics* 69). If we take Manovich's claim that data is also a kind of language seriously, then it is obvious to compare language with data, it is possible to draw parallels between digital and non-digital hermeneutics. However, this idea clearly raises a problem from the point of view of defining information and interpreting it in a momentary way. For language, like man, exists in an analogue world and does not function in the same way as data.

Manovich's ideas on cultural analytics and artificial intelligence aesthetics extend this debate further by introducing methods that use large-scale data analysis and artificial intelligence to understand cultural trends. Manovich argues that the digital turn represents a move towards "meta-media", where traditional media types are not discarded but recycled and transformed through digital manipulation. This process results in new forms of artistic and cultural expression that challenge our previous ideas about media and aesthetics ("New Media" 23)

Byung-Chul Han is a Korean-German philosopher and theorist who critiques the information society primarily from an ideological perspective, and his writings focus on the interconnections between technology, capitalism and culture. His general thesis is that the pervasive influence of digital media is fundamentally reshaping human relationships, creating an environment where constant connectivity leads to anxiety and superficial interactions. In his most resonant work, The Burnout Society, Han traces how neoliberalism is reshaping work and personal life through an analysis of how it emphasizes self-optimization and performance, often leading to burnout and mental exhaustion.

Foucault, like Kittler, is an important thinker for Han, often a point of reference in his work. The starting point for the concept of "infocracy" is that instead of the disciplining society of Foucault, we now live in a society of transparency, where the subject voluntarily deprives himself of freedom. As Han puts it, "isolation is replaced by interconnectedness, while in digitality communication becomes surveillance" (Infocracy 10), ultimately pointing to the totalitarian tendencies of digitality.

At this point, it is worth mentioning Shoshana Zuboff's concept of observation capitalism. Zuboff defines it as a kind of capitalist logic of production, a wildcat of the capitalist establishment. Also, part of this definition is that the machinery of surveillance capitalism is an "anti-democratic and anti-egalitarian [...] top-down, market-driven coup" (Zuboff 479). The author writes of this epistemic coup in a similar way to Han on infocracy: "By annexing human

experience, this coup leads to an exclusive concentration of knowledge and power that gives privileged influence over the distribution of knowledge in society: the privatization of the most important principle of the social order in the twenty-first century" (479).

Zuboff, however, makes a clear distinction between observation capitalism as an operation and digitality as a technology: "Observation capitalism is not a technology; it is a logic that permeates technology and commands it to act. Observation capitalism is a form of market that is unthinkable outside the digital milieu, but it is not identical with the 'digital'" (21). However, it seems to me to underestimate the totalitarian tendencies of digitality as a Kittlerian discursive network. Zuboff goes on to argue that the denial of this distinction is in fact in the vested interest of the surveillance capitalist regime itself to deflect blame: "That surveillance capitalism is an operational logic and not a technology is a vital statement, because surveillance capitalists want us to believe that their practices are inevitable manifestations of the technologies they employ." (21).

Yet we cannot ignore the fact that certain technologies facilitate certain cultural patterns. Just as McLuhan has shown the social changes linked to book printing (individualism, rationalism, etc.), digitality also attracts tendencies towards transparency, observation, concentration of knowledge. Thus, even if we cannot say that digitalization is observation, it is not an exaggeration to say that, by virtue of its transparency, digitalization becomes the priority of observation. Of particular note is the phenomenon that, while digitality creates a society of transparency, it helps one to a more complete understanding and in no way leads to a more democratic episteme. Han convincingly demonstrates these characteristics of an information society based on digitality.

An important related term is "psychopolitics," which refers to the subtle but pervasive control exercised through psychological manipulation and self-exploitation (*Psychopolitics* 11-12), primarily in the interests of what Zuboff calls the "Big Other" of surveillance capitalist tech companies. Han clearly contrasts this notion with the traditional conditioning of disciplinary power: "It [the voluntarily submissive] is the subject's self-creation as product and performative" (*Infocracy* 8). Beyond the ideological issues of knowledge and power, there is a market logic to Han's formulation of psychopolitics. In a media environment where we can theoretically create and access an infinite amount of information, we are left with an inexhaustible limited resource. In Shackell's words, "if cognition in the information age has re-entered our field of vision as a limited resource whose mode of allocation, though

misunderstood, is no longer inviolable, then a theory of semiosis based on finite cognition will be essential to understanding the world to come" (240).

And finally, we have to ask: what is the role of the humanities in all this? We seem to talk a lot about digitality, digitalization, we seem to recognize the seriousness of the phenomenon, but we leave it to other disciplines or tech companies. This seems to me to be a missed opportunity. It is time to finally take digitality and the infosphere seriously and treat them as legitimate objects of study. If we turn our backs on culture, digitality will devour us along with humanity.

2.4 Conclusion

To conclude this chapter, I attempted to outline the foundations of my dissertation, establishing the theoretical model necessary to analyze digital cultural phenomena. This chapter posits that the processes of digitalization and datafication should be understood as the ruling episteme, termed digitality, following a Foucauldian approach. The goal of constructing this framework is to provide a context for interpreting intermedial cultural and epistemological phenomena inherent in digitality. The chapter started with an overview of the impact of the digital turn in the humanities, marking a shift from purely analogue practices (such as close reading) toward quantitative methods and Big Data approaches. This transformation led to important reflections on epistemology, questioning how algorithms and data shape what is considered knowledge and the sustained importance of hermeneutics to keep pace with computation. I heavily relied on Kittler's concept of "Discourse Networks" to model relations of culture and technology, arguing that the rise of digitality represents a fundamental background to how technology and institutions select, store, and process data.

Furthermore, I highlighted the determinism/constructivism debate concerning technology and culture, resolving this dichotomy through Stiegler's model of inherent human technicity and the concept of epiphylogenesis, which treats technology not as an external tool but as a fundamental condition of human existence. A key focus of this chapter was the epistemological challenge presented by the data-centric environment, which is increasingly described as an "anti-hermeneutic landscape". I argued that the humanities, particularly hermeneutics, are necessary for redefining and repositioning human interpretation within this environment.

Finally, I would like to quote the French philosopher and mathematician Bruno Bachimont, who outlines the real stakes of digitalization by outlining two possibilities:

> The first [possibility] says that human beings are by nature semiotic animals that individuate meaning depending on the dynamics of contexts and situations they live in. Therefore, interpretive freedom is inherent to human behavior. But, and this is the second possible answer, our technical systems are increasingly globalizing and interpersonalizing communication, and hence are normalizing it according to planned situations that are provided by the technological communication systems themselves. Therefore, the invention of everyday life and ordinary situations is being progressively reduced to being no more than

one of the possible combinations that have already been constructed on the basis of the formatted data (30).

The following chapters of my dissertation will reflect on the possibility of Bachimont's second answer and aim to explain why the first answer should be our answer.

3. Digitality as a Discourse Network: from Atoms to Bits

As we have seen in the previous chapter, in Kittler's media theory the discourse network comprises technical media, symbolic systems and practices that together define the limits and possibilities of knowledge. In 1800 it was print and literature, while in 1900 it was the new analogue media. The logic of digitality as a network of discourse is fundamentally based on discrete layers of digital data. A philosophy of digitality must therefore reveal how digital systems are built on discrete, operational layers of data and how these determine what action or meaning digital media can produce. This chapter explores this thesis, drawing on theorists who arrive at the primacy of the underlying data layer from different perspectives: the media materialism of Friedrich Kittler, the information philosophy of Luciano Floridi, Alberto Romele's concept of data as data model. From these perspectives, a common insight emerges: the digital way of being is shaped by what can be recorded, stored and processed as data. In the digital world, only that which can be transformed into data and expressed in bits that can be connected, stored and processed is truly existent. All the layers or other phenomena that run on top of these can ultimately be created on the basis of this data infrastructure.

In the discourse network of digitality, the position of knowledge and culture is determined by digital media: computers, networks and ultimately binary code. What we can say, see or remember is inextricably intertwined with the functions and limitations of these technical systems. As Kittler later put it, "What remains of people is what the media can store and transmit" (*Gramophone* XL). In the digital age, this aphorism suggests that our identities and memories are effectively reduced to and survive in the data we produce. If something cannot be stored as data, it cannot be preserved within the framework of digital existence. This strict media materialist stance raises an ontological implication: in the digital world, to be is to store or process something as data. The opposite is epistemological: what cannot be recorded as digital data is outside the horizon of formal knowledge or experience. Kittler's theory of media thus shows at first glance why the data layer must be seen as primary — because it is only through the media (today, digital data media) that anything becomes part of the worldly discourse.

Kittler's famous statement that "software does not exist" sums up the extreme materialism of his approach ("There Is No Software" 223). By this, Kittler suggests that software is ultimately an illusion — an interface that obscures the reality that all computation

is reducible to hardware operations. Underneath the user-friendly abstractions of applications and graphical user interfaces, the computer is constantly manipulating the flow of binary signals. As Kittler explains, "all code operations, despite their metaphorical capabilities, such as "call" or "return" are ultimately reducible to entirely local string manipulations, and are ... notations of voltage differences" (223). Every program, image, or text in a digital system is essentially encoded in machine-readable binary states. If we attempt to strip away every layer of programming languages and interfaces, we come to the conclusion that there is no software at all: instead, a series of bits printed on physical media. This insistence has deep ontological significance: it means that the essence of digitality lies in the discrete states of a material medium. Epistemologically, what we know as digital, or what we interact with in digital form, always remains within the boundaries of this mechanical layer. However rich a multimedia experience may seem, "we simply do not know what our writing is doing" at the machine level (221); yet it is this level, the level of the data, that performs and enables the experience. Kittler's media materialism, therefore, forces us to recognize digital data as the invisible reality of our technological world: a vast, functioning text composed of ones and zeros, a new discourse network in which human meanings exist only by virtue of their encoding in the underlying data structure.

While media materialists and cultural technology theorists draw attention to the physical and historical foundations of digital data, philosopher Luciano Floridi takes the debate into the realm of ontology. According to Floridi, a fundamental re-ontologization is taking place, whereby reality is increasingly understood in terms of information ("A Look into…" 60). He introduces the concept of the infosphere to describe the ubiquitous digital information environment in which we live today (Floridi, *Fourth Revolution* 41). As the boundaries between the online and offline worlds disappear, people and digital entities become part of a single information space. Floridi observes that modern information and communication technologies (ICTs) have become "environmental forces" — like a digital ecosystem — that shape every aspect of our lives. We are seamlessly interconnected and effectively integrated into the infosphere, to the extent that we are increasingly living an "onlife," a term Floridi uses to describe the blurring of boundaries between physical and virtual existence.

In this infosphere, existence is indistinguishable from accessibility as information. People, objects, transactions — all leave digital footprints, and all appear and act as data. Floridi characterizes people as "inforgs", organisms of information, to emphasize that our identities and interactions are now intimately intertwined with information processes (*Fourth Revolution*

94). The notion of the infosphere thus extends the notion of digitality as a discourse network: it is not only media that define our situation, but our situation is a vast information medium in which everything that exists is in some form of data. Ontologically, Floridi's philosophy of information conceives of information as the stuff of reality. It brings into coherence and provides a systematic framework for ideas such as John Wheeler's famous phrase "It from bit," according to which physical reality can be understood fundamentally as information (*Fourth Revolution* 70).

Floridi's argument, particularly in *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, is that our age is forcing us to rethink the fundamental questions in the light of ICT: who are we? What is reality? Just as Copernicus, Darwin and Freud pushed humanity off the center of the cosmos, life and reason, so the digital revolution is a "fourth revolution" that pushes us off the center of the infosphere. No longer privileged observers outside the world of information, we are part of it and defined by it. This view leads to the ontological principle of digitality: anything that cannot be transformed into information — that cannot be accessed informationally — is at risk of being ignored or considered non-existent. Conversely, if something can be represented as data, it can enter the processes that govern reality — it can be analyzed, reproduced, disseminated and used.

Floridi's infosphere is thus another way of describing the universal discourse network of digitality, in which discrete information is the common currency of existence. From an epistemological point of view, Floridi's observation that our knowledge of the world is now largely filtered through information filters: search engines, databases, data analysis, is important. The structure of the data we capture and the way we process it shapes the structure of what we know. The concept of the infosphere therefore suggests that understanding reality requires understanding information structures. Floridi's perspective thus confirms the primacy of the data layer, since it virtually identifies it with reality. If Kittler showed that there is virtually "no software, only hardware," Floridi suggests in a more metaphysical tone that there is no reality, only information. For Floridi, digitality is an ontological condition in which the infosphere — the set of available data and information processes — is the new nature in which we exist.

Where Floridi sees a seamless infosphere, Alberto Romele stresses the lack of continuity and the need for interpretation. Following the tradition of philosophical hermeneutics, Romele argues that data are essentially traces — traces or remnants of reality that require interpretation to make sense of them. In his book *Digital Hermeneutics*, Romele proposes a shift from

thinking in terms of information to thinking in terms of traces and data, highlighting the raw registry nature of digital records. Romele describes the process by which digital traces are transformed into data and data into information (*Digital Hermeneutics* 38). This formulation reminds us that what we call information in a digital context is not given in advance; it is the result of an active process: events leave traces, these traces are collected as data and only through analysis or contextualization do they become information (usable knowledge). Romele's sequence of trace → data → information underlines the epistemological point that data is central as a key axis. Data are not fully meaningful in themselves — raw registrations of traces — but they are no longer the original event or phenomenon — their processed, formatted version. They function as indexical, transformed, delocalized references to something in the real world.

Drawing on hermeneutic philosophy, and in particular Paul Ricoeur's concept of distanciation, Romele argues that the defining characteristic of the digital experience is the distanciation between things and their becoming data (*Digital Hermeneutics* 101). We increasingly know the world through its data traces, and we must interpret these traces in order to reconstruct their meaning. According to Romele, digitality thus has a double significance: on the one hand, it gives us an unprecedented amount of traces (diaries, records, data sets), but on the other hand, to understand it, we need to follow these traces and acknowledge their mediating role.

Romele advocates what he calls "a hermeneutic alternative to the concept of semantic information" (*Digital Hermeneutics* 75). Rather than assuming that information is transparent and immediately interpretable, Romele wants us to pay attention to recording and registering as fundamental operations. He writes, for example, "I think that today recording, registering and tracking [are]... at the heart of our digital practices" (*Digital Hermeneutics* 72), suggesting that these activities form "a difficult but possible matrix of epistemology" (*Digital Hermeneutics* 77). By this he means that in the digital age our knowledge is based on tracing — a difficult epistemology because traces are fragmentary and require interpretation, but possible because we have the means to trace and analyze them. As an example, take web analytics: millions of clicks and views (traces of user behavior) are recorded as data; interpreting these (information) requires interpretation, pattern recognition, inferring user intentions and so on. Raw data is meaningless in itself until hermeneutic work is done on it. Romele's perspective thus complements Floridi's: yes, we live in an infosphere, but it must be read as a text full of signs and clues, not just calculated.

Ontologically, Romele does not dare to say that reality is information. In the digital age, reality is accessible through traces and records, and these are always partial. There remains a gap between reality and its description. Following Bruno Latour's insights, our reality has become, according to Romele, "a global laboratory in which entities and events can be followed step by step" (*Digital Hermeneutics* 54) — everything is recorded and logged — but "reality and the virtual never coincide". The map is not the territory, even if the map is updated in real time with infinite data points. A hermeneutics of digitality therefore pays attention to what is lost or hidden when we reduce things to data, even as it acknowledges that data is our primary means of relating to things. Epistemologically, Romele's focus on the trace reminds us that digital knowledge is indirect. It is always knowledge acquired through a representation — a data trace.

While there are many historical precedents for this, the scale and degree of mediation in digital culture is unprecedented. We know ourselves through our data (health indicators, social media profiles), others through their digital footprints, and society through datasets and opinion polls. Romele points to the danger of taking these traces as our entire reality. The hermeneutic approach insists on interpretation and context: data should be interpreted as evidence of something, not as things themselves. Romele thus returns to the importance of the layer — not to glorify it, as Floridi does, but to examine it with a critical eye. She asks us to acknowledge the strengths and limitations of the discrete data layer: its strengths lie in the fact that it allows for tracking and analysis, but its limitations lie in the fact that it abstracts and simplifies the phenomena it tracks.

When the idea of Romele's trace as data is considered in conjunction with the ideas of other theorists already discussed, a fuller picture of digitality as a discourse network emerges. Despite their differing emphases, these perspectives converge on one key point: the discrete, operational data record is primary in digitality. Whether we approach it from the perspective of hardware and code, historical markup techniques, philosophical ontology, or hermeneutic interpretation, we recognize that digital existence depends on what is stored and processed as data.

What cannot be stored or processed becomes epistemologically mute or ontologically powerless in the digital world. Digital culture is therefore a culture of recordings and computations, of circulating data. To exist ontologically in the digital age is to leave data traces or to be encoded in a database. Epistemologically, what we know is increasingly based on the manipulation of these traces: computational results, database queries, the interpretation of log

files. The primacy of discrete data does not mean that human meaning or culture is reduced to zeros and ones, but that all meaning-making is now based on the mediation of discrete symbolic operations. In this light, I will analyze the concept of digital data in the remainder of this chapter.

3.1 The Concept and Ontology of Digital Data

By the 21st century, digital data has become a dominant feature of life, permeating virtually every area from science and culture to governance and everyday social interactions. As we navigate what is often described as the era of "Big Data," understanding what data is and how it operates within cultural and epistemological frameworks is of paramount importance. This session will explore the theoretical foundations and cultural interpretations of digital data from the perspectives of philosophy of technology, media studies and cultural studies.

In its most basic sense, the term "data" derives from the Latin datum, meaning "that which is given" (18). This origin refers to something presented as a factual given, and indeed, the early use of "data" in mathematics and theology referred to accepted givens or premises for the sake of argument (19). However, the apparent simplicity of "given things" conceals a complex range of interpretations around the concept of data. Floridi, in a 2008 article published in the *International Encyclopedia of the Social Sciences*, identifies several levels of meaning of the term data: an epistemic interpretation (data as collections of facts or evidence), an informational interpretation (data as information content) and a computational interpretation (data as binary symbols or signals in computers) ("Data" 2-3). Each of these captures a part of the meaning of data. For example, in the epistemic sense, scientists talk about experimental observations as "data" that underpin knowledge claims, and in the everyday sense, data are seen as facts that we collect. In the informational sense, we often equate data with information — as in terms such as "personal data", which refers to information about an individual. And in computational terms, all content — text, images, music — stored or processed by digital devices is encoded as binary data (ones and zeros).

Yet, neither interpretation alone fully captures the essence of the data. Floridi therefore proposes a more basic definition: the diaphoretic interpretation, which defines data as an elementary difference — "x is different from y" ("Data" 5). In this conception, data in its most abstract form is a relational entity, a pure distinction or lack of unity in the world ("Data" 5). For example, even a blank page can be considered data when contrasted with a page containing text — the difference (the presence or absence of signs) is itself informative ("Data" 7). This is consistent with the idea that data need not carry an inherent meaning; it can simply be a sign or a signal that can be interpreted. By defining data as distinctions, we can emphasize the raw potential of information — distinctions that can be detected and measured ("Data" 9).

However, whether the data actually convey meaning or truth depends on interpretation and context. This is where a framework informed by the humanities can offer a counter perspective to the understanding of data as an objective given. Theorist Lisa Gitelman reminds us that data is anything but "raw" and cautions against thinking of data as a natural resource to be mined. Rather, data should be seen as a cultural product to be produced and interpreted. In Gitelman's terms, "raw data is an oxymoron", meaning that data is always "cooked" by the methodology and context in which it is collected (3). Zsolt Almási also warns in connection with Shakespeare's production history in Hungary that "preconditions [for data selection] must be clarified in order to make the analysis possible at all" (147). In other words, every dataset involves choices about what to collect and how to represent it; data does not simply fall from the sky, but is selected, measured and coded within a human framework.

This humanities-based approach challenges the notion of data as neutral pieces of information. Instead, data are seen as inherently laden with theories and values, and their meaning is only revealed through interpretation. Even in technical contexts, this is recognized in definitions of data. The philosopher of science Sabina Leonelli also defines "data" as a relational category — any research result (observation, measurement, image, etc.) can be considered data if and because it serves as evidence in support of a claim in the context of a particular study (817). Data, therefore, "in and of themselves have no truth value and cannot be considered as simple representations of particular phenomena" (811). Instead, data are understood as part of an inferential process: they are mobilized, packaged and communicated as evidence for their "prospective utility" in support of knowledge claims (811). This perspective emphasizes that what counts as data depends on context and purpose — a set of numbers or records becomes "data" the moment it becomes the basis for an argument or reasoning.

Historically, the concept of "data" has undergone a major transformation, evolving from a concept in rhetoric and philosophy to a central element of modern science and digital culture. In a brilliant study, Daniel Rosenberg explores the concept of data, its etymology and the philological history of the term. In the early use of the term, especially in the Enlightenment and pre-Enlightenment periods, the term data refers to things given in an argument — often truths given at the beginning of an argument (for example, axioms of geometry or principles of theology) (19). The historian of science Daniel Rosenberg notes that "from the beginning, data was a rhetorical concept," bound up with the context of the argument (18). Since data meant

what was given before the argument, its meaning was inherently malleable, changing with different "argumentative strategies and contexts" (20). As the practice of inquiry has changed over time, so has the understanding of what constitutes data. The rise of modern natural and social sciences in the 18th century created new conditions of argument and new assumptions about facts and evidence. However, the pre-existing semantic structure of the term "data" gave it important flexibility in these changing circumstances (21). Scientists began to use "data" to describe empirical observations of the world — measurements and recordings that could serve as objective evidence. Crucially, Rosenberg stresses that the term's prior semantic flexibility gave it a useful adaptability: even as the nature of facts changed with experimental science, calling something "data" carried an aura of legitimacy as a factual basis for knowledge (21).

By the 19th and 20th centuries, particularly with the advent of statistics and large-scale data collection (e.g. government censuses, scientific laboratories), data were increasingly seen as epistemic currency — a basic unit of knowledge that, when aggregated and analyzed, produced insights about nature and society. The development of computing in the mid-20th century then dramatically accelerated this trajectory. The Second World War and the Cold War ushered in an era of so-called Big Science, characterized by huge data sets (for example in nuclear physics or socio-economic planning) and the computers needed to process them. As Gitelman and Jackson put it, if the mid-20th century helped usher in Big Science, "the new millennium has arrived as the era of Big Data" (2). Digital technologies have made it possible to collect, store and transmit data on a scale previously unimaginable — every click, every movement potentially counting for something, to someone, somewhere (3).

This explosion in the sheer volume and variety of data represents, in Gitelman's words, "a seismic shift in the way data is understood and used today" (2). Data is no longer confined to scientific research or government bureaucracy; it has become a ubiquitous by-product of everyday life and a key resource for industries. People are starting to talk about a "data revolution", where data itself is a driver of innovation, economic value and social change. Data has become a culturally fashionable buzzword — with promise and power attached to it (think "data-driven decision making" or "data is the new oil"). At the same time, critical voices have begun to remind us of the lesson of the past: even in this new guise, data do not speak for themselves. Gitelman, Almási and others have provided a timely reminder that no matter how large or automated our data collection, "data are always already cooked and never completely

raw" (*Raw data* 3). In other words, today's vast datasets are still being produced by someone, for some purpose, using specific technologies and methodologies.

The historical trajectory of data can thus be seen as a journey from scarcity to abundance, from curated data (observations carefully recorded in a laboratory notebook) to the ubiquity of datafication (the continuous recording of human activities through digital services). Along this trajectory, the status of data as epistemological objects has also changed. In earlier periods, data were often secondary to theory — one collected data to test or illustrate a hypothesis. In the modern era, data sometimes precedes theory: huge data sets are collected without a specific hypothesis in mind, in the belief that patterns will emerge inductively. This reversal has fueled debates about whether data now play a new role in knowledge production or whether it is simply a new scale of an old practice.

Importantly, the global dimension of this story cannot be ignored. While much of the discourse around the evolution of data is rooted in the Western scientific and technological tradition (and English-language sources), the data revolution is a global phenomenon. Concepts such as open data, digital repositories and data analytics have become internationally pervasive and part of global scientific and political culture. However, the historical understanding of data remains deeply influenced by early intellectual traditions that conceptualized data as objective facts. The work of scholars such as Gitelman, Drucker and Rosenberg serves as a corrective, adding a cultural and historical consciousness to our understanding of the data. They remind us that the idea of data is already formed — and can be reconstituted. As the following chapters will discuss, this has important implications for the philosophy and ethics of data in contemporary society.

What kind of thing is data? This ontological question is also a concern for philosophers and theorists of the information age. There are several key debates about whether data are fundamentally objective entities that exist independently of observers or constructs that depend on interpretive frameworks. There is also debate about how to distinguish data from related concepts such as information and knowledge. This section explores the different philosophical perspectives on the nature of digital data and the debates they have inspired.

One of the classic ways to place data is within the data-information-knowledge hierarchy. According to a well-known formulation, often called the DIKW pyramid — although its origins are obscure and methodologically unsound — data are raw symbols or observations,

information is processed or interpreted data that has meaning, and knowledge is assimilated information interpreted in context. More useful is Rosenberg's typology, which reflects rather on categories of cognition: "facts are ontological, evidence is epistemological, and data are rhetorical" (18). They perform fundamentally different, not necessarily sequential, functions in the process of knowledge production.

Although hierarchical approaches can provide a useful shortcut, we have long discussed nuances. Floridi, for example, defines information as "well-formed, meaningful and true data", suggesting that data only becomes information when it is endowed with semantic content and correctness (366). This suggests that data can be incomplete in itself — a set of numbers or characters that need context to mean something. However, not all data becomes pure information. As Floridi points out, there can be sets of data that are semantically meaningless, but are still data. The relationship between data and information is thus not a simple equivalence; some data represent information about a referent (e.g., a weather log shows temperature over time), while other data (e.g., an encoded image file) may require an interpretive key to reveal information about 'something" (354-355). Data, especially digital data, may exist in a state of potential meaning that requires interpretation to unlock.

The ontological debate often focuses on whether data are the product of interpretation before or always before interpretation. The realist view tends to treat data as pre-existing facts to be discovered. According to this view, digital data generated by binoculars or a social media platform are recordings of reality — they may contain errors or noise, but they are supposed to reflect something "out there" in the world. The constructivist or interpretivist view, on the other hand, argues that data is inseparable from the apparatus (conceptual, social, technical) that creates it. Johanna Drucker's argument that the notion of data should be rethought as capta emphasizes this: what we call data is in fact the result of observation and collection practices, not intact inputs from the natural world (no pag.) Here we can draw parallels with the debates about observation in the philosophy of science: just as observers are laden with theories (no observation is free from the influence of the observer's expectations, theories and tools), so data are laden with theories and tools. In this respect, data do not exist as pure, theory-independent entities. They are ontologically relational — a point echoed by Leonelli's definition of data as "relational categories" defined by how they function as evidence in a given context (817).

Another philosophical question is whether data can be considered objects in themselves, or properties, or relational categories. Floridi's diaphoric definition (data = difference) implies

that data are not material objects, but rather abstract — they exist as distinctions that can appear on different physical media (magnetic dots on a hard disk, ink on paper, electrical signals, etc.). This begs the question: is a digital data (say the number 42 stored in a computer) identical to its physical embodiment (the particular configuration of electrons or magnetic domains that represent the number 42)? Many information philosophers argue that data are independent of any particular material form, since the same data can be copied or transmitted on different media without losing their identity. The digital character of modern data reinforces this idea that data are discrete, immaterial entities that can be infinitely replicated. However, media theorists caution against treating data as completely immaterial: the material infrastructure that enables digital data (servers, networks, storage devices) is crucial to the existence of digital data and determines what data can do (for example, the speed of the network affects what real-time data applications can be implemented).

The ontology of data is also intertwined with questions of truth and representation. It can be tempting to think of data as representing reality in a straightforwardly true or false way. Rosenberg's historical analysis, however, reminds us that calling something "data" "makes no assumptions about its veracity" (37). Data can be inaccurate, biased or even deliberately falsified and still be data (37). A set of numbers may be entirely fictitious, yet function as data within an analysis (although it may lead to false conclusions). This non-dependence on truth value is partly why the concept of data has proved so useful: we can collect data without immediately having to verify their truthfulness, deferring this issue to later analysis or debate. In Floridi's formulation, truthfulness is a property of the information or knowledge, not of the raw data itself ("Is Semantic Information" 365-366). Data are simply symbols or signs in themselves; whether they accurately represent something in the world is a separate question. Thus, according to one philosophical view, data are semantically open: they carry no truth value unless they are interpreted in a propositional context.

The larger philosophical debates also address the question: is data discovered or created data? In the context of digital data, this question is: do sensors and algorithms discover data that was lying dormant in reality, or do they create new representations that did not exist before? Critical theorists often argue for the latter: data about human behavior, for example, are often produced by imposing formal categories on activities (for example, the malleable nuances of social relationships are transformed into binary categories of "friend" or "not friend" on a platform). The term "datafication", which I will discuss in more detail later, captures this idea

of transforming aspects of life into data. Ontologically, this approach means that data is not simply lying in the world waiting to be picked up, but is created using instruments, sensors and classification systems. In their paper on classification, Geoffrey Bowker and Susan Leigh Star show how the way information is recorded (say, a medical diagnosis schema or an ethnic category in a census) determines what data can exist and be collected (47-48). Data in this view are performative: they create certain realities by making some things countable and others invisible.

This data-driven epistemology is often referred to as a kind of inductivism or "data-driven paradigm." Rather than knowledge leading from theory to data (deduction) or from data to theory (classical induction), knowledge in the Big Data paradigm is often claimed to arise from the data itself. Fields such as genomics, astronomy and particle physics were early adopters of this approach, sifting through petabytes of data after new discoveries — sometimes finding phenomena (such as new genomic correlations or celestial objects) in the data first, before formulating a theory. The social sciences have followed suit, with computational social sciences mining social media data for patterns of human communication or movement that no theory had predicted.

But the initial enthusiasm was tempered by critical realizations. One major criticism is that correlation is not causation. Mass correlations revealed by algorithms can be spurious or insignificant. Given enough data, random patterns can always be found; distinguishing meaningful signals from noise becomes a challenge. Statisticians have been warned about the problem of multiple comparisons — if millions of hypotheses are tested (as data mining algorithms implicitly do), some will only appear significant by luck. The "end of theory" view underestimates the role of theory in guiding which contexts to trust and which to follow. Anderson's vision, for example, may find that Google searches for ice cream correlate with stock market performance, but without theory it may falsely proclaim predictive insights that are merely coincidental.

Another criticism focuses on bias and context. As boyd and Crawford noted, big data sets are not a perfect representation of the world; they are created in a specific context and often contain systemic biases (2). They warn that "claims of objectivity and accuracy are misleading" — big data analyses may create the illusion of objectivity, but the selection of data sources, the cleaning process and the algorithms used all convey subjectivity (4). Social media data, for example, may appear to be a direct window on public opinion, but platforms have specific

demographic characteristics and algorithms that shape what people see and do (7). Without understanding these, our "data-driven" findings may be distorted. This has given rise to debates about data quality versus data quantity: more data is not automatically better if the data is biased or irrelevant.

The era of big data has seen the emergence of machine learning models (in particular deep learning) that are remarkably accurate in their predictions, but often opaque — the so-called "black box" models. This leads to an epistemic question: do we have knowledge if we can predict something correctly, but cannot explain why or how the prediction is made? Traditional philosophy of science would say that understanding requires explanation, not just prediction. Some data scientists argue that in practice, prediction is sufficient for many tasks (e.g. a machine learning model can predict equipment failure without explaining it, yet it can be extremely useful). On the other hand, we could argue that this is a form of "knowledge debt": by foregoing understanding, we may be unable to correct mistakes or apply knowledge to new scenarios. The drive for interpretable AI, as opposed to black box models is partly about reconciling robust pattern recognition of big data systems with the human need for interpretability and causal insight (Rudin 206).

There is also a social epistemological dimension: who can know and how? Big data often requires expensive infrastructure and expertise, leading to a divide between those who have access to the data (and the tools to analyze it) and those who do not. boyd and Crawford refer to this as the "Big Data rich and the Big Data poor" (13) divide. Knowledge production can be monopolized in the hands of large institutions (large technology firms, well-funded laboratories) that can afford to collect and process big data, potentially marginalizing other forms of knowledge creation (such as small-scale field research or indigenous knowledge systems).

In response to these problems, some scholars advocate mixed methods and "middle-way" epistemologies. Rather than being seen as a substitute for traditional methods, big data approaches can be seen as a complement to them. For example, the notion of abduction (inference to the best explanation) becomes relevant: patterns found in the data (induction) can suggest hypotheses that are then tested or explained by theory (deduction). This cyclical process (sometimes called the data-theory loop) has arguably always been the way science has developed, but big data accelerates the inductive phase. An illustrative case is computational social science: an algorithm may detect an unusual set of activity in social media data (inductive

discovery), then social scientists may qualitatively investigate it to form a hypothesis about what social phenomenon caused it (explanatory theory), then test that hypothesis through another data set or experiment (should be a source)

Another adaptation was the emergence of Critical Data Science or Critical Algorithm Studies, which injects reflexivity into data science workflows. Here, practitioners question their own models and findings, check for biases, and incorporate domain knowledge at every step. Rather than "numbers speak for themselves", the ethos is that "numbers need context to speak meaningfully". This is consistent with the humanities perspective discussed earlier, ensuring that data analysis is guided by theoretical and ethical reasoning.

In some areas we see epistemological adjustments. In biology, Sabina Leonelli notes that big data practices have led to data being treated as a portable commodity that can be reused across contexts, challenging the idea that data only have meaning in their original experimental context (816). The new challenge is to document data so well (metadata, provenance information) that it can be meaningfully reinterpreted elsewhere — in effect, it takes context with it. This is interesting from an epistemological point of view, because it recognizes that data are not facts in their own right — they need context to be useful as evidence, especially when they are placed in a new context.

In summary, big data epistemologies promised a radical change in knowledge production, and while they did indeed transform many practices, they did not make theory and human understanding obsolete. Instead, they have encouraged us to re-examine how we derive meaning from empirical analysis. The current understanding is more sober and nuanced: large-scale data can reveal patterns of unprecedented scale and granularity, but human insight is needed to verify, contextualize and explain these patterns. As one commentator has put it, in the age of big data, "more is not simply more — more is different" (Anderson). We have more data, but that changes the questions we ask and the ways we need to answer them responsibly. The goal is not to choose between data-driven and theory-driven approaches, but to integrate them.

This has profound consequences for knowledge production. It suggests that future researchers will need to be skilled in both statistical/algorithmic techniques and critical, theoretical thinking. It also highlights the importance of interdisciplinary collaboration: data scientists, domain experts and ethicists working together to ensure that knowledge from big

data is not only reliable in the narrow sense of accuracy, but meaningful, relevant and aligned with a broader understanding of the world. Ultimately, humanities-oriented frameworks — with their emphasis on interpretation, context and critique — are not the enemy of big data, but an indispensable partner in turning massive amounts of data into real knowledge.

Given the complexity of the topic under review, the need for interdisciplinary engagement seems undeniable. Digital data is at the intersection of technology and the humanities. Philosophy of technology contributes insights into the informational nature of reality (e.g. Floridi's infosphere and the idea of the fourth revolution, where the boundary between humans and information is blurred). Media and cultural studies contribute to understanding how data circulate in society, how narratives about data (e.g. the promise of Big Data or fear of surveillance) influence public thinking, and how cultural biases can be encoded into technical systems. The humanities have an important task to do with data being part of a centuries-long search for knowledge and meaning — and thus bringing to the table centuries-old ideas about meaning, ethics and epistemology that remain highly relevant. As a result, the study of digital data is inherently a collaborative effort across disciplines.

3.2 Data as Performative and Afformative

Another approach to the production of knowledge starts from an examination of the nature of speech acts. When we say "knowledge," we often mean some kind of linguistic statement (e.g. a scientific proposition, a historical statement, the result of a literary analysis). How do these statements become valid? What distinguishes a mere opinion from what we take to be accepted knowledge? This question can be answered by J.L. Austin's speech act theory and Jacques Derrida's reflections on it.

In the 1950s, J. L. Austin developed the theory of speech acts, which is based on the idea that language use not only describes reality, but also performs an action. Austin pointed out that there are so-called performatives: statements which, when uttered, also perform something. A classic example is that when I say "I do" (or "I do") in a wedding ceremony, I am in fact entering into a marriage bond — the linguistic act is itself an action (32). Austin contrasted performative statements with constatives, which are simply true-false statements: the statement "it is raining" is either true or false. Performatives, on the other hand, are "neither true nor false", but rather can be either successful (succeed) or unsuccessful (fail), depending on whether the appropriate appropriateness (or "happiness") conditions are met. These conditions relate, among other things, to the appropriate social context: for example, for a wedding "yes" to be valid, the presence of the right people (registrar or priest, the parties to be married), the observance of the right procedure, etc. If these are absent — say, if the "yes" is pronounced in the context of a joke or a play — the performative act remains "empty" or unhappy (infelicitous) (40).

Austin's insight is important because it shows that the power of linguistic utterances lies not in the sentences themselves, but in the context and conventions that make them valid. Even a simple promise ("I promise that...") is only a real act (promise) if the speaker means it and the hearer accepts it — otherwise it remains an empty word. In the same way, a scientific statement (e.g. "Water boils at 100°C under normal atmospheric pressure") becomes accepted knowledge under certain conditions: if it is experimentally supported, if it is communicated according to the publication norms of the scientific community, if it is repeated and confirmed by others, etc. In other words, statements that communicate knowledge have their own "conditions of appropriateness" — institutional, methodological and linguistic conditions.

Although Austin was primarily concerned with oral manifestations, he also extended his thought to written performative acts. For example, a written contract or a will also has performative force: if it is drawn up in the right form and in the right circumstances, it produces legal effects (transfers of property, succession, etc.). The specificity of verbality, according to Austin, lies in the fact that it is carried out in direct interaction, face-to-face, and thus provides a framework for feedback and immediate interpretation. Writing, on the other hand, is more autonomous — we will return to this later in relation to Derrida.

In his essay "Signature, Event, Context" Derrida examines Austin's theory and the concept of communication in general. Derrida agrees that linguistic meaning cannot be fixed by reference to the speaker's intention and the immediate situation alone. Indeed, he goes further: for him, the existence of any linguistic sign, whether spoken or written, presupposes a degree of contextual independence and reproducibility, which he calls iterability. Derrida points out that when we speak or write something, the sign can be detached from its original context and reused in another situation, for another purpose, with another meaning (Kell source) This detachment does not entail the sign becoming unintelligible — on the contrary, it allows the sign (the text) to persist in time and space and to be communicated with again and again. A written text, for example, remains intelligible in the absence of its author, and even centuries later, it continues to have an impact in other cultural contexts. According to Derrida, this autonomy and iterability is not only a property of writing, but a general property of language — it is also present in speech, just not as evident.

Derrida is critical of Austin's assumption that performatives in a theatrical or fictional context are to be treated as "empty," as "not having happened" ("Signature" 16). Austin says, for example, that an oath or promise spoken on stage is not a real oath, but a pretense — he labels such use as "parasitic" use, which he says is outside the scope of his theory ("Signature" 17). Derrida, on the other hand, points out that it is precisely the possibility that a linguistic act can be quoted, repeated in a different context (even in fiction), that shows the essential property of language: that meaning can never be completely closed off by a given intention and context. Inherent in all meaning is the possibility of 'slippage" in meaning, of misunderstanding or reinterpretation ("Signature" 18). Indeed, according to Derrida, the possibility of linguistic communication itself is created by the constant transgression of signs beyond their original position — if this were not the case, language would be no more than a series of one-off manifestations that could not be recognized or recalled.

Derrida thus argues that contextual openness and iterability are a prerequisite for language — and thus for the communication of knowledge ("Signature" 17). The implication is that in communication there is never a complete guarantee of meaning: there is always room for maneuver, the message is never one hundred per cent controlled by the author or the circumstances. Meaning is not closed, but always open to future reinterpretation. If it were not, language would simply not be able to function as a language.

In this sense, I would argue when Derrida speaks of the instability of meaning, he is pointing out that knowledge is performative in nature: in order for a statement to become knowledge, something must be done with it — applied, interpreted, repeated, confirmed. Scientific facts also become "facts" because they are accepted and used by a community through a network of experiments, publications, references. But this process always involves the possibility of reinterpretation. Derrida is in fact extending Austin's notion of performativity: not only are linguistic actions such as an oath or a promise performative, but every statement is performative in some way, insofar as it has an effect and is incorporated into a discourse.

The performativity of knowledge production also means that we decide what counts as valid knowledge in different ways at different ages, the condition of flourishing or "appropriateness" changes. In the Enlightenment, for example, what counted as legitimate knowledge was that which was published and quotable in proper print, with an author, publisher, and censorship, such as the Encyclopedia. In contrast, in the digital age, the criteria for the verification and authenticity of information face new challenges (e.g. establishing the authenticity of an online source, issues of plagiarism and AI-generated content, etc.). Just as, according to Austin, the success of a performative act depends on context, so the "truth" of knowledge depends — in a sense — on the conditions provided by the cultural-technical context, rather than on a transcendent standard. This does not mean that there are no natural laws or objective reality — only that what we treat as knowledge is always a consensus performed by a community and mediated by technical media.

To declare a sequence of bits as data is in itself a speech act in the Austinian sense: an inscription which, by its very naming, evokes a normative horizon in which the object can circulate as evidence. "Raw" bits do not have epistemic status in themselves; they become data only when a competent actor places them in an abstract but socially recognized framework — file formats, ontologies, timestamps, institutional archives — that corresponds to what we might call "conditions of usability." These conditions are not empirical truth-values, but pragmatic:

readability for future analysts, aggregability and acceptability by the relevant epistemic community. As the title of Gitelman's already cited volume ‚*Raw data is an oxymoron* reminds us, even the simplest CSV is full of such conventions, proving that knowledge does not begin with observation, but with "this counts as data" by performative attribution.

If we further apply Derrida's insight on iterability, we find that every successful data assignment also provides infinite iterability for the file. Once placed in the infosphere — described, following Floridi, as the ambient medium for all information entities — the dataset can be indefinitely separated from its original intent, replicated, remixed and placed in new contexts. Each replica remains formally identical (for example, an SHA-256 hash), but acquires new illocutionary power in the hands of new interlocutors; what is today a population statistic may tomorrow serve as training input for a neural network. The characteristics listed by Kitchin as defining Big Data — "volume, speed, variety, completeness" (1) — can therefore be understood rather as a modality of this iterability, a characteristic of a signal system whose capacity for repetition transcends any authorial or institutional control.

Consequently, digital data embody performativity in the Derridean sense: they can only function as data if they meet the criteria of community flourishing, but this authority can be maintained precisely because data is structurally open to future, unforeseen re-enactments. Epistemic stability and semantic slippage are thus interrelated: the more rigorously a data is formulated for use, the more contexts it evokes, and the less its meaning can be definitively fixed. In the digital abstract, knowledge is thus neither a mirror of reality nor a mere text game, but a continuous bargaining between the Austinian force that binds a data set to rule-governed practices and the Derridean drift that lets the same data set escape into new, as yet uncharted acts of knowledge.

Werner Hamacher takes Derrida's deconstructive critique of Austin's binary oppositionist, so to speak structuralist, system even further, and introduces what he calls the affirmative dimension of language. Hamacher writes at greater length about the term in an extremely long footnote in his essay "Affirmative, Strike": he uses affirmative to designate a fundamental linguistic force that underlies, indeed precedes, all the performatives (1143). The term is a deliberate counterpoint to the term "performative," suggesting something like a formative force that never fully emerges (1139). As Hamacher puts it, the affirmative is "the enabler that cannot be fulfilled in any form," an event of forming that remains formless — "as both enabler and disabler, as action and non-action at once" (1139). Most importantly, it is not

merely a negation of form or a failed performative; it is "not affirmative, not a negation of the formative": rather, the affirmative is the latent potentiality in any speech act that enables it without ever being itself realized as a bound act or positive content (1139).

In Austinian terms, we could say that the affirmative is a force of language that precedes all conventional rules of illocution. It is the "pre-capability" of meaningful utterance, which is not realized in a performative, which "does not strive for realization" in the conventional sense (1139). In Benjamin's sense, it is analogous to the pure medium of communication: a means that serves no external purpose but simply enables communication to take place. Benjamin outlined a "politics of pure mediation" in which means are "pure" precisely because they "do not serve as means for ends outside the sphere of mediation" (1139). Hamacher applies this logic to language. Just as Benjamin's general strike is a political action that suspends all instrumental ends (it does not enact or enforce law, but disrupts the whole system of means-ends relations), so the affirmative linguistic act suspends everyday reference and illocutionary success. It "merely allows, but never posits" (1139) — that is, it allows meaning or effects to emerge while never stabilizing them in a definitive form or outcome. All performative speech acts, even the most definitive, are thus forever indebted to this affirmative "letting". Hamacher writes that even the absolute performative (say, a sovereign legislative utterance) "must first of all be subjected to the absolute other of deposition — affirmation" (1139). In other words, before any speech act can successfully position something (a truth, an action, a social bond), it is already permeated by the movement of de-position ("deposition"), which renders it ungrounded. The affirmative is this constant de-positioning force — an indeterminate potential in language that enables performative acts but also empties them of their certainty.

Hamacher's concept of affirmativity thus names a linguistic grounding potentiality that is never realized as a determinate speech act. It is "foundational but not realizable": a condition of the possibility of meaningful utterance, which nevertheless does not itself produce a tangible speech event or illocutionary result. If Austin's performative is directed towards producing effects by virtue of certain conventions, Hamacher's affirmative resides in the moment before utterance is condensed into any conventionally recognizable act. This is a liminal force, an act of opening (as the Latin ad- in the affirmative suggests) without closure (1139). Hamacher describes this as "the mere possibility of language as such," a layer that "always allows language to happen" but never as a finalized presence (1139).

In his essay “Lingua Amissa,” he even describes language itself as nothing more than an unfulfilled potential: “Language is nothing other than this unfulfilled, unrealizable promise of language” (153). In other words, language speaks as a promise — a potential meaning or action to come — which by its very nature is never fully fulfilled. This affirmative promise is not a concrete message, but the very fact that language holds itself in reserve, always more than any particular instance of speech.

Hamacher’s affirmative conception is in line with broader deconstructive readings, in particular Derrida’s ideas of iterability and the “messianistic” structure of promise. According to Derrida, outlined in an essay titled “Signature, event, context”, any sign with meaning must in principle be repeatable: iterable, which means that it can break out of any fixed context and generate new, unforeseen significations. This is not an accident of language, but the very essence of language. It follows that every utterance is to some extent “quotable” and contingent. A well-done speech act cannot exclude itself from quotation or variation; indeed, its effectiveness derives from a code or convention that can be quoted by anyone.

Every utterance implicitly promises its own meaning by referring to common norms, but this promise is never entirely under the control of the speaker. It is an open-ended promise of a context and future reception that is yet to come. In Derrida’s words, “the performative of this promise is not one speech among others. It is implied by every other performative, and this promise proclaims the uniqueness of a language to come” (“Monolinguailism” 67). Here Derrida, like Hamacher, identifies an immanent promise in language: a commitment to meaning and effect that accompanies and transcends even the most banal statement. This ‘structure of promise... informs all speech” as an irreducible condition (“Monolinguailism” 21).

Turning to the field of digital data, we find a striking structural affinity with this affirmative philosophy of language. Digital data — binary code sets, information databases — form the basis of our digital systems today. On the surface, data may appear purely instrumental: machine-readable bits designed to be processed to achieve specific results. However, a theoretical reflection reveals that data, like language, operates through deferral, openness and a kind of non-actualization that parallels affirmative structure.

Repeatability built into digital data at the most basic level. A digital file can be copied, pasted, transmitted and recontextualized indefinitely without losing its form. One code can run in countless environments; one data set can feed countless algorithms. This repeatability is

analogous to Derrida's signal iterability. It means that digital objects inherently transcend any specific context of use. For example, a text file or a binary-encoded image can be run on different platforms or inserted into different documents — just as a word or phrase can be quoted in new settings to achieve different effects. The meaning or function of data is therefore not fixed or self-contained. Just as the power of a sentence depends on how it is received by the listener (or even by future readers beyond the author's intention), the meaning of a piece of data is only revealed in the process of execution or interpretation, which can always change. In this sense, data carries afformative potential: a raw capacity that can be realized in multiple ways, but is not identical with any one realization.

Consider the ontology of a dataset in a database. It is itself a sequence of bits — not a "single purpose". It can encode a number, a name, an image, or nothing meaningful at all until a program parses it out. This is similar to Hamacher's view of language as a potential, "linguistic promise" that is not yet real. Data can also best be thought of as a promise: they promise meaning or utility (why else would they be collected and stored?), but this meaning is deferred until the moment of use, and even then is not final. When a program queries the database and uses the data, it realizes one of the latent possibilities — but the data can always be queried again in a new way, or used by another application for a different purpose. The original data therefore goes beyond any particular use. No single application exhausts the meaning of the data. In practical terms, this is why the same dataset can support countless analyses, or why code libraries can be reused in unexpected projects. From a theoretical point of view, it indicates that data as such is a non-functional reserve: it retains its full essence and remains available to be reused.

This withholding is analogous to the afformative "merely letting" that Hamacher attributes to language. Data "allows" information to be produced without itself being a determinate message. In its purest sense it is a medium, a carrier of possible meanings. We might say that digital data is only "data at all, in terms of a future" use or interpretation, echoing Hamacher's assertion that language is only language in terms of a future language. Until it is put into a particular software or interpretive context, data is in a state of suspended utility — much like an unread text or an unspoken thought. And even once it has been used, it can be put back into a state of suspended utility and await another use. The structure here is undeniably one of suspension and deferral rather than immediate presence. The power of data lies in what it can enable, not in what it does. In this way, digital data also embodies a kind of messianic

time: always oriented towards a future. We often talk about the "potential" of data (in terms of insights, innovation) — a notion eerily close to the messianic promise that is never fully revealed. Just as Derrida's messianism is an expectation of an event that never appears (a truth or truth to come), so too data is imbued with an expectation of future meaning that no single outcome can satisfy. There is always a "next" algorithm or a revised interpretation on the horizon.

Moreover, instrumentalizing data — using it for practical purposes — always leaves a residue that is not instrumentalized. No matter how tightly software is written to use data in a fixed way, the data can usually be re-exported, recombined or simply read in a different way. In other words, data is resistant to being completely consumed or used by any program. This is analogous to Hamacher's affirmative dimension, in which any particular use of language is insisted upon as surplus or residue that cannot be integrated. Similarly, in digital systems we find that the meaning of data is never totalized by direct application. There is a residual reuse. Even attempts to delete or erase data often leave traces (backups, logs) that can be restored, suggesting that data is inherently intended to persist rather than be used once.

From a philosophical point of view, we can say that digital data has an "ultra-transcendental" status, not unlike Hamacher's description of affirmative language. It is beyond existence in the sense that it is not a condition of a concrete existent (not an object with a concrete meaning), but of many existents with many meanings. It is transontological -it is prior to concrete ontic entities such as "this picture" or "that document", since it can become any picture or document with the right code. This makes data a kind of pure representational tool. In fact, the digital realm formalizes what Hamacher sees in language: an infinite play of means without ends, without final signifiers. A line of code, for example, can generate a result (an end), but this result can always be treated as new data — a new means — for a further computation. The process is open-ended, iterative, and postpones the final output. This is why, in contemporary discourse, data is often likened to a renewable resource or wellspring — its value lies not in the finished product, but in its continued usability.

By looking at digital data through an affirmative lens, we emphasize that data are not static, fixed things, but dynamic possibilities. Like language, it can be rewritten and reread. And it is affirmative because it suspends realization: data on a hard drive is physically real, but does not correspond to any realized meaning until it is activated by a software or application. Even then, the "meaning" (e.g. a visual representation, a semantic output) is temporary — digital data

thus resists becoming a complete instrument — it always retains an inoperative side. One could say that in digital culture, as in Hamacher's philosophy, the medium is never completely dominated by its messages. There is always a gap between what data is (at an electrical or binary level) and what it means or does, and in this gap lies the realm of indeterminacy and freedom.

To conclude, Hamacher's affortmative concept provides a rich theoretical framework for understanding both language and digital data as structures of deferred realization. According to this conception, language is not a means of simply carrying out intentions or conveying stable meanings, but a field of virtual acts — promises, potential significations, gestural openings — that do not solidify into a final form. Digital data can also be seen as affortmative : a substrate that holds open countless possibilities for computation and meaning-making without ever collapsing into a single final outcome. In both cases, we are dealing with a game of possibilities and reality, means and ends, iteration and presence.

This affortmative play suspends the fulfillment of meaning and action, yet it is what makes language and digital data so generative. It is the surplus, the "not yet", that is the source of the creative force. By theorizing data through affortmativity, we resist the temptation to treat data as a mere tool or a finished information commodity. Instead, we value it as a fundamentally poetic element — a creation that is never fully finished — akin to the infinite web of language that weaves a future "language" to come. Such a perspective reinforces the realization that in the digital age, as in the world of texts, meaning and effect are never simply given, but always in the making , always hovering in the affortmative interval between what is coded and what is yet to be realized.

3.3 Digital and Data Hermeneutics

The proliferation of digital traces, complex algorithms and interactive interfaces has not eliminated the need for human understanding, but on the contrary, has made the hermeneutical task even more critical. Since data alone does not carry a clear meaning, the act of interpretation becomes crucial. Thus, hermeneutics, the theory and methodology of interpretation traditionally applied to texts, has been extended to the field of digital data and media. The question arises: how do we "read" data? One of the key questions today is how to apply traditional interpretive principles when the "texts" in question are data sets, computer interfaces and algorithmic outputs. Leading digital culture scholars, such as Alberto Romele and Paolo Gerbaudo, argue that hermeneutic approaches are not only applicable to data, but are in fact necessary to fully understand the meaning of data in our socio-cultural context, as I will attempt to demonstrate throughout this chapter.

Traditionally, hermeneutics starts from the premise that any object of study (a text, a work of art, an action) only makes sense through an active process of understanding that takes into account context, intention, and the interaction between the parts and the whole. The discourse of the digital age often falls into the temptation to treat data as self-interpreting facts, or to assume that algorithms objectively reveal the truth in numbers. Gerbaudo, Romele and others, however, remind us that digital data are not self-explanatory givens, but constructs that are interpreted from the outset. Romele builds on this insight, highlighting the inherently interpretative nature of all digital artefacts. The interface of an application or platform presents information within a certain framework, privileging some elements while hiding others, just as a text emphasizes certain themes over others. Similarly, the outputs of algorithms (such as the classifications of a machine learning model or the results of a search engine) encode the assumptions and training data of their creators. Without hermeneutics, we risk taking these outputs at face value and ignoring the layers of meaning already built into them. In short, the proliferation of digital data requires not less, but more hermeneutics to make sense of the flood of information.

Data hermeneutics is an emerging term for this digital interpretive approach. As Gerbaudo describes it, quantitative data analysis needs to be complemented by "data hermeneutics", which focuses on interpreting the deep meaning structures of social media conversations and other digital contexts (96). In practice, data hermeneutics can involve what

Gerbaudo calls “close reading of data” (97), the adaptation of literary close reading techniques to digital content. For example, in analyzing social media data on a political protest, a data analytic approach can produce statistics on trending keywords, network graphs of user interactions or sentiment ratings. These are valuable, but they may lack context — the irony in a hashtag, the cultural references in a meme, the narrative arc through a series of posts. In a hermeneutic approach, researchers would focus on specific conversations, interpret the pattern of tweets or posts qualitatively, and decode the symbolic and rhetorical nuances (106). Such an approach treats social media posts as texts to be interpreted, rather than just data points to be counted. A researcher might discover, for example, that a seemingly divisive hashtag is used sarcastically by one group and seriously by another — a layered meaning that quantitative analysis alone can ignore.

Gerbaudo notes that while big data methods have provided sophisticated ways to measure and visualize online dynamics, “data analysis tends to ignore their content and the deep meaning behind the dynamics” (99). After an initial period of naïve enthusiasm, many have come to realize that quantitative metrics can be misleading out of context. In the same paper, Gerbaudo, quoting Crawford and boyd (217), points out that it is wrong to think that “bigger data is always better data” or that numbers alone guarantee objectivity (99). He therefore argues for a reintegration of interpretive, qualitative methodologies — essentially an updating of hermeneutic traditions from fields such as phenomenology, literary criticism and cultural anthropology — to address the specific properties of digital data (96). This “continued relevance” of interpretive approaches is seen as a necessary counterweight to the algorithmic analysis that dominates digital policy research (99).

The work of Alberto Romele represents a significant extension of the hermeneutical tradition in the face of the challenges and specificities of digital media and technology. Rather than simply applying Gadamer’s or Ricoeur’s theories to new examples, Romele actively innovates and adapts hermeneutics for the digital age. In her Digital Hermeneutics (2019), Romele aligns herself with Ricoeur’s philosophical hermeneutics, importing Ricoeur’s notion of distanciation — the idea that understanding requires a certain distance or difference between the interpreter and the object of interpretation. For Ricoeur, distanciation explains how texts acquire meaning beyond the authorial intention and invite the reader into an open process of interpretation. Romele extends this idea to digital objects, noting that digital technologies create both distance and proximity in interpretation. For example, a set of data abstracted from lived

experience creates distance (people's behavior is replaced by numbers), yet the same data can be interactively visualized or algorithmically analyzed to bring patterns "closer" to our understanding. Romele transforms Ricoeur's detachment and gives Ricoeur a material and technological aspect beyond his primarily linguistic focus. In doing so, he opposes what he (following Latour) calls the "materialist idealism" of Gadamer, Heidegger and Ricoeur — the tendency of traditional hermeneutics to privilege language as the exclusive medium of meaning — and insists that hermeneutics must also take into account non-linguistic media, from images to interfaces, as vehicles of meaning (2). By expanding the scope of hermeneutics to include digital forms of expression, Romele ensures that interpretive theory speaks to websites, social media postings, databases, and other media central to contemporary digital culture.

Romele's key contribution is the notion of "digital hermeneutics as material hermeneutics." He argues that the interpretation of digital information cannot ignore the material and technical infrastructure that underpins it. In contrast to a purely conceptual or textual analysis, Romele's approach looks at how hardware, software, code and design are involved in meaning-making. He defines his hermeneutics of digital information as an explicitly material hermeneutics for three main reasons: first, it "starts from an analysis that is internal and not independent of the object in question", meaning that the digital object (an application, a database, an algorithm) must be treated on its own terms, and its internal logic and structure must be understood rather than a predetermined interpretive scheme (Romele 38). Second, it is "concerned with the diversity of contexts in which meaning is produced and received" (ibid.). In the digital space, this means looking at how data is produced (by whom, by what means, on the basis of what assumptions) and how it is received or used by audiences or users. A tweet, for example, only makes sense if you take into account the platform's capabilities, the author's intent, the algorithmic timeline that displays it, and the readers who interpret it. Thirdly, digital hermeneutics is "interested in the material — the techniques and technologies through which digital traces are transformed into data and data into information" (ibid.). Here Romele emphasizes that each step of the process — from the recording of digital traces (such as server logs or user clicks) to their structuring as data and finally their representation as "information" — is mediated by technical means and human decisions. By focusing on these material aspects, Romele extends hermeneutics beyond classical textual interpretation to a comprehensive analysis of the building blocks of digital culture. Romele's digital hermeneutics thus bridges the gap between philosophical interpretation and the practical, often technical, work of data management, which is particularly important in digital humanities research.

Another way of extending Romele's hermeneutics is the notion of mimesis (or emplotment) in digital contexts. Adapting Ricoeur's theory of narrative, Romele sees digital data analysis as following a kind of narrative arc: data does not speak for itself but is configured into a story or pattern by the algorithms and models that process it. When, for example, large amounts of social media data are analyzed to identify trends, or when algorithms generate user profiles from browsing history, these processes resemble the construction of a narrative — the selection and combination of elements into a coherent structure. According to Romele, databases and algorithms perform what Kant called categories, or what narrative theory calls "emplotment": they bring order to the raw flow of events (Romele 107). In doing so, they also exercise what Romele, drawing on Kant and Ricoeur, calls productive imagination. Modern algorithms, in particular those powered by machine learning, are able to recognize patterns and draw inferences on a scale and complexity beyond direct human capabilities. This can be seen as a form of imagination, where the machine "connects the dots" in creative or non-obvious ways. Romele, however, does not imply that the machine's imagination replaces the human interpreter. On the contrary, he is exploring how these algorithmic interpretations require a second-order hermeneutics: we humans must interpret the interpretations produced by machines. For example, if an artificial intelligence model produces profiles of individuals as part of an "algorithmic governance" (insurance or police profiling), we need to interpret and critique these profiles and the criteria behind them. Romele's work thus highlights the two-layered hermeneutic process of digital culture: algorithms interpret data, and we have to interpret algorithms and their results. This insight extends the hermeneutic philosophy of to the fields of artificial intelligence and data science, suggesting that understanding today often means understanding how our tools generate meaning and what assumptions they embed. Romele's approach extends hermeneutics by considering this new interpretive horizon where human and machine action intersect.

Finally, the digital traces — the records and traces left by human activities in the digital environment — are at the heart of Romele's hermeneutical argument. In contemporary discourse, digital traces (such as browsing histories, social media posts, GPS locations) are often treated as a rich source of insight into human behavior. They feed everything from targeted advertising to scientific research on social trends. Romele approaches digital traces from a hermeneutical mindset rooted in philosophy. In Paul Ricoeur's *Memory, History, Forgetting*, he draws on his discussion of traces, where he interprets a trace as a sign of something that was, pointing back to a missing cause or event. In the digital case, a trace (say

a set of timestamps from smartphone usage) is a sign of a person's past actions. The critical question is how to get from the trace to the meaning. According to Romele, we should never treat digital traces as self-evident facts, but rather consider them as clues to be interpreted, much like a historian considers archival documents or an investigator considers evidence in the field. He cautions against two extreme attitudes: on the one hand, the positivist belief that digital tracing allows us to see society exactly "as it is" (the illusion of complete transparency), and on the other, the skeptical view that human life is infinitely irreducible to any data (the view that digital representations are practically useless or misleading). The concept of tracing strikes a balance. Romele suggests that when talking about data as traces, we should emphasize the origin of data from lived reality and its transformation through recording. As one of Romele's commentators summarizes, referring to his insight, "to speak of data as traces rather than information is to recognize the mimetic operation from which data derive and the ontological gap between the reality from which they derive and the virtual they create" (D'Alessandris 174-75). In other words, when we call a data set a collection of traces, we acknowledge that it has a referent (something that has happened), but also that it is a representation shaped by means and interpretations. Romele's digital hermeneutics insists that to understand the data we need to understand this path from reality to record to analysis. In this way, Romele extends hermeneutics to treat digital traces in a similar way to how we treat historical documents: as partial, interpretable evidence to be contextualized, rather than as a full portrait of reality (Digital Hermeneutics 75). This approach fundamentally enriches the interpretation of data: it brings humility and depth to data analysis, reminding analysts that data points are hints and clues to stories, not complete stories themselves.

In many areas, influenced by big data and algorithmic analytics, there was an implicit hope that, given enough data and computing power, the numbers would simply speak for themselves. Romele's work challenges this perception by showing that hermeneutics provides the critical human context without which data remains meaningless or misleadingly simplistic. The extension of hermeneutics to the digital humanities serves as a corrective to approaches that rely solely on computation. In digital literary studies or cultural analytics, for example, algorithms can be used to discover patterns in thousands of texts or tweets. While such methods may find correlations or trends, Romele would stress that interpreting these patterns still requires humanistic insight: patterns need to be linked back to cultural context, authorial intent, genre conventions or social circumstances in order to make sense of them. In a 2020 study, Romele and colleagues illustrated this by examining how Twitter data can be used to gauge

political opinions, applying Ricoeur's hermeneutic circle (or triple mimesis) model to interpret the results (Romele, Severo and Furia 80-84). They have shown that one's conceptual framework (prior knowledge) influences what data is collected and how it is analyzed, that methods and algorithms (configuration) impose their own structure on the results, and that the researcher must then interpret the results in light of the original questions and the broader context (reconfiguration). This iterative process is fundamentally hermeneutic: it recognizes that analysis is not a one-off objective reading, but a dialogue between expectation and evidence, between the part and the whole. By explicitly mapping this process, Romele offers a methodology for digital humanities that keeps interpretation at its core, ensuring that quantitative analysis does not take place in a vacuum.

Romele's hermeneutic emphasis also brings to the surface the ethical and existential dimensions of data interpretation that quantitative methods alone may ignore. If we see algorithms and data as offering interpretations of us (profiles, risk scores, personalization), then hermeneutic engagement means questioning and understanding these interpretations. This has profound implications: it reminds us that human freedom and identity cannot be fully captured by an algorithmic profile or data set. As Romele argues, human beings are not fully transparent either to themselves or to others; there is always an element of "hetero-determination" — shaped by external forces — but there is also an element of unpredictability and self-interpretation (Romele 27-28, 148). Hermeneutics, especially in the tradition of Gadamer and Ricoeur, has long argued that understanding ourselves is the infinite task of interpretation. Romele extends this to the digital self: our digital footprints provide new mirrors in which we see ourselves, but these mirrors are imperfect and require interpretation and critique. Thus, Romele's work enriches the interpretation of data by adding layers of reflection: what does it mean that an algorithm has predicted this about me? What narrative does my data feed into? How does the interface I use shape my perception of reality? These are hermeneutic questions that ensure that data analysis remains connected to human meaning and values. In practice, this may mean using qualitative analysis alongside quantitative analysis, involving interdisciplinary perspectives (a philosopher or anthropologist working alongside data scientists), or developing a 'sociological imagination" adapted by Romele to the digital age. The result is an approach to the digital humanities and data science that refuses to reduce understanding to computation alone. Instead, it integrates technical analysis with interpretive depth, enabling empirically grounded and philosophically informed insights.

Ultimately, a hermeneutic and humanities approach to data insists that meaning and interpretation are central, even when dealing with large-scale, numerical or algorithmic information. This resists any assumption that data 'speak for themselves". Instead, it argues that we talk to the data: it is through the dialogue between the data and our knowledge of the world that we arrive at understanding. This approach does not contradict computer analysis but rather enriches it. As explained above, the best results often come from a synthesis — for example, using data mining to find an unexpected pattern, and then using interpretive analysis to explain that pattern (or to check that it is something real and not an artifact). In the context of knowledge production, this reflects the interplay between quantitative and qualitative methods and highlights why the humanities are indispensable in the age of data. They ensure that the human element — values, intentions, interpretations — is not lost in the digital sea of information.

3.4 Big Data and Critical Data Studies

While hermeneutics is concerned with the interpretation of the meaning of data, Critical Data Studies (CDS) extends the subject to power, politics and ideologies surrounding data. In recent years, scholars in media studies, sociology and cultural studies have explored how the widespread "datafication" of society is reshaping social relations, often in problematic ways. This chapter discusses the following key critical concepts: datafication, data capitalism, data mazes and the sociotechnical shaping of data.

Datafication refers to the process by which aspects of the world and our lives become data. In contemporary usage, it refers to "the transformation of human life and social interactions into quantified data" (Mejias & Couldry, 2). Social media platforms, for example, date our friendships (turning interactions into likes, shares, number of followers), our preferences (through clicks and views), and even our movements (through location). José van Dijck describes datafication as part of a larger 'social media logic" and notes that the use of data as a tool to understand and track people's behavior is becoming a leading principle across sectors (202). Van Dijck argues that datafication is rooted in certain ontological and epistemological claims — namely, the belief that social phenomena can be fully captured in data, and that the analysis of such data will lead to objective, actionable knowledge (202). This belief system is called dataism, a "widespread secular belief" in the power of data to accurately guide society (201). Dataism is so successful, van Dijck argues, because people, often unwittingly, trust their personal data to corporate platforms and extend that trust to other institutions that manage data (such as research or law enforcement), dataism essentially equates more data with more truth, reflecting a kind of technological utopia or determinism about data (202).

Critical scholars point out that datafication and dataism carry hidden assumptions and implications. The first is the assumption of neutrality — the idea that data is a reflection of reality. In practice, data is quantified in terms of what can be measured and what is considered valuable to measure. This can lead to biases. For example, if police activity is quantified using crime statistics and predictive policing algorithms, this can reinforce existing biases (excessive policing of certain neighborhoods will result in more data on crime there, which justifies additional policing — a feedback loop). Another consequence is the erosion of privacy and autonomy: as van Dijck notes, metadata and personal data have become "regular currency" that

citizens pay, often unknowingly, in exchange for digital services ("Datafication" 197). The ideological aspect of datafication is that it normalizes pervasive surveillance and data collection as not only inevitable but also desirable for progress or security. Van Dijck calls for a more critical look at this "whole ecosystem of connected media", pointing out that government, business and even academia can collude (consciously or not) in promoting the ideology of datafication ("Datafication" 204).

Closely related to datafication is the concept of data capitalism. Data capitalism refers to a form or stage of capitalism in which data is a key resource and commodity — the primary source of monetization and value creation. In the digital economy, many services are "free" for users, but profits are made by extracting and exploiting data. Social media platforms, search engines and apps collect behavioral data and monetize it through targeted advertising or by selling insights. Sarah Myers West for example defines data capitalism as a system in which the commodification of personal data allows for an "asymmetric redistribution of power" to the benefit of those who control the data (4). Some technology companies accumulate huge data sets, which become barriers to entry and sources of algorithmic advantage, thereby concentrating power. Shoshana Zuboff's concept of surveillance capitalism further illuminates this dynamic: users' experiences are observed (tracked through data) and transformed into products (predictive data about behavior) that are sold or used to manipulate behavior. In this model, the user is effectively both a product and a consumer — their data is sold and also targeted with data-driven content.

Data capitalism, critics argue, has socio-political implications: it can undermine individual agency (through ubiquitous surveillance and the influencing of behavior), challenge existing regulatory frameworks (as data flows often bypass traditional definitions of property or commerce), and even reshape class structures (some talk of a new class of "data workers" — people whose digital work produces data, often without remuneration). The term data colonialism was introduced by Nick Couldry and Ulises Mejías to draw parallels with historical colonialism. They argue that the expropriation of human life through the continuous acquisition of data is a new form of colonialism: "the subjection of human beings to new kinds of relationships built around the extraction of data" (4). Just as colonial powers expropriated land and resources, today's digital corporations are appropriating not only personal data, but also the very social activity and communication through which that data is generated. This perspective

broadens the critique of data capitalism by highlighting issues of sovereignty, inequality and resistance at the global level.

Another key theme of CDS is the imaginary of data — the collective ideas and narratives about what data is and what it can do. These ideas shape policy, innovation and public opinion. For example, the imagination of big data as omniscient has led to initiatives that pour resources into data collection, based on the assumption that more data inherently solves complex problems (in governance, health, security, etc.). We see "smart city" initiatives based on the idea that if enough data is collected on traffic, energy use and citizen movement, cities can be optimized to near perfection. Likewise, the notion of data as truth fuels things like the Quantified Self movement, through the belief that self-tracking data can reveal truths about an individual's health or behavior that subjective experience might overlook, as illustrated for example in a study by Pols et al. (104-105).

However, critical studies show that perceptions of data often mask important realities. Take the case of policing, for example: the notion of "data-driven policing" promises unbiased enforcement through analytics, but in reality, can reinforce biases and create a false aura of objectivity around discriminatory practices. Or consider the "open data" imaginary, which assumes that making government data transparent automatically empowers citizens — it may or may not empower already powerful entities (such as data-mining companies) if citizens lack the tools to harness the data. Critical scholars urge us to examine whose imagination serves: positive narratives about data often come from those in power who benefit, while marginalized groups may have a different imagination (for example, if data is seen as a tool for surveillance and control).

3.5 Conclusion

To conclude, I would like to reiterate the reason critical data analysis emphasizes the socio-technical shaping of data. This means recognizing that data are produced by intertwined social and technical processes. Every dataset has a social life: people decide what to collect, how to label it, what database schema to use; later, people decide how to interpret and act on the data. Similarly, technical choices (what sensor to use, what format to store the data in, what algorithm to use) enable and constrain what the data can tell us.

Medical data is a very clear example of this: if a health database does not have a category for a particular symptom because it was designed without the consent of a particular patient population, then that symptom essentially "does not exist" in the data system, which can bias research and treatment. Data therefore reflect the values and biases of their creators and maintainers. Critical studies often rely on Science and Technology Studies (STS) to analyze how standards, classifications and infrastructures (such as cloud platforms or APIs) shape the potential of data. They also look at the politics of the data infrastructure: who owns the servers, who writes the code, what laws govern cross-border data flows?

By making the socio-technical nature of data visible, we can better address issues of accountability and ethics. For example, if a data set used in a machine learning model leads to a discriminatory result, rather than blaming "the data" in the abstract, a sociotechnical approach would trace the result back to specific collection methods, historical biases, and design decisions, and assign accountability accordingly.

Lastly, it is important to stress that critical data analysis is a necessary counterweight to uncritical data enthusiasm as exemplified several times in the chapter. It reminds us that data, far from being purely neutral inputs, are deeply intertwined with social contexts and power structures. Datafication changes the way we live and understand the world, but its supposed objectivity is often rich with ideological biases. Data capitalism creates wealth and innovation for some actors, while also championing new inequalities and forms of serious exploitation. Data fantasies about the bright future of the field can inspire progress, but they can also mask and justify harm.

The sociotechnical perspective that is becoming more prevalent in the scholarship requires us to examine human decisions at all stages of the data life cycle. Together, these

insights make it possible to move us towards a more critical data literacy that asks who benefits, who suffers harm, and what assumptions are at play when we encounter a data-driven initiative. This critical perspective is essential to ensure that data practices are consistent with democratic values, equity and human well-being.

4. Digital Data Operations as Cultural Techniques

After exploring the concept of digital data and its interpretative framework, I will focus on the culture-shaping role of the discourse network of digitality. As we have seen, in 21st century digital culture, data is not merely a passive resource, but an active shaper of knowledge, culture and possibly subjectivity. The pervasive datafication of digital life — the transformation of behaviors, objects and ideas into quantitative digital information — has profound cultural implications (Mejias & Couldry). In this chapter, I argue for an approach to digital data operations as a cultural technique in its own right, rather than as a neutral by-product of technology. In German media theory, Kulturtechniken (cultural techniques) refers to a set of practices, skills and tools that operationalize basic distinctions and thereby create cultural order and meaning (14). Using this framework, I consider digital data not as a mere technical output, but as a central object of theoretical inquiry: an active process that underpins 21st century knowledge systems and symbolic practices. This approach allows for a humanities-oriented examination of the technical underpinnings of digitality and to ask relevant questions for knowledge production. I consider the description of digital data as a cultural technique as the nova of my dissertation.

In this chapter, I draw on the insights of German media theorists as interpretative frameworks, rather than isolated overviews: in addition to Kittler, already mentioned several times, the works of Bernhard Siegert, Sybille Krämer, Cornelia Vismann and Markus Krajewski allow us to explore the operational logic, epistemological foundations and symbolic or representational dimensions of digital data in culture. Kittler's media materialism, for example, insists that culture itself derives from the ability to select, store and process relevant data through networks of technologies and institutions (*Discourse networks* 369). Siegert's approach to cultural technologies "dissolves the notion of media into a network of operations" that reproduce and process the distinctions that are fundamental to a given culture (*Cultural Techniques* 13). Similarly, Krämer's work on the history of digitization situates digital data within a long continuum of symbolic practices (the alphabet, numerical notation) that flatten and discretize reality ("Flattening" 11). In *Files*, Vismann shows through the genealogy of the file that record-keeping techniques have continuously mediated power and knowledge, which reappear in computer architectures as files, folders and registers. And Krajewski's media history of catalogues and servers reveals proto-computer data infrastructures that long predate the

digital age. Weaving together these perspectives, we can understand procedures of digital data as an active cultural operation: one that not only rewrites the world in bits, but also transforms how we know and interpret that world

While Kittler emphasizes the electronic hardware and binary code that form the basis of the digital network, other German media theorists approach the primacy of the data layer through the concept of Kulturtechniken, or cultural techniques. In particular, Bernhard Siegert and Sybille Krämer examine the fundamental operations — the actions of recording, computing, sorting — that create the cultural order. Their work shows that long before the emergence of electronic computers, culture was already formed by discrete, algorithmic operations on symbols and material elements. Digitality is the culmination of a wider historical process of transforming reality into manageable, countable units. Siegert provides fascinating examples of how cultural techniques discretize the world. He talks about the emergence of the zero as a positional marker in 16th century printing, and how a scientist's introduction of dots to mark missing text "transformed real gaps into a series of discrete, countable elements. Reality is digitized" (Siegert, monograph, 27). What was once an indescribable absence becomes a countable zero, an element of a series. By digitizing the absence, early modern typographic techniques made the text machine-like: iterable, copyable, and subject to formal operations. According to Siegert's history of media, cultural techniques ranging from writing systems to surveying break down the continuum into classifiable and processable discrete units (27). This is precisely the logic that applies to digital data. Building on Siegert, digitality can be seen as the ultimate cultural technique: a technique that discretizes all inputs and thus renders them operative. Here, however, we emphasize that data are the basis and the possibility of these techniques; in the digital age, entities have acquired an autonomous, infrastructural existence.

Sybille Krämer also highlights the role of operational signals in the history leading to digital computing. She sees modern computation as the result of the "emergence of operational symbolism" (Krämer 93) — symbols are not only used to represent ideas, but also to perform operations. In both mathematical writing and computing, symbols are tools that can be manipulated independently of semantic meaning, much like bits that can be flipped syntactically in computers. Krämer points out that the separation of symbols from their semantic context was a key cultural technique that paved the way for computers (90-95). Practices dating back to Leibniz effectively pre-digitize reality by creating discrete symbol systems that can be systematically processed. Krämer's recognition that the ability of machines to process data depends on their prior abstraction of reality into discrete, state-like symbols. Thus, in the

genealogy of digitality, we find a continuum from the cultural techniques of writing and computation to the machine operations of computers. Discretion and rule-bound manipulation are key. When we engage with digital systems, we interact with operational signals — bits, codes — that have no meaning in themselves, only what we attribute to them at a higher level. The digital layer of data follows its own logic, and this is a prerequisite for any higher level meaning to be realized on a machine.

Markus Krajewski's historical studies further demonstrate that the primacy of operational data is not a new feature of late 20th century computers, but part of a longer infrastructural genealogy. Krajewski's media archaeology research on card catalogues reveals that by the 19th century, information was already managed in a discretized, modular way similar to digital databases. In his book *Paper Machines*, he calls the library card catalogue a "paper machine" and argues that card catalogues "have all the basic logical elements of universal discrete machines: they store, process and transmit data" (3). Each card is a discrete, mobile unit (similar to bytes or database entries today) that can be sorted and retrieved according to strict rules. The information contained on a single card can be rearranged infinitely — it is a "discrete, uniform and mobile medium" (3) that has enabled new knowledge management practices. According to Krajewski, the card catalogue system was in principle an early Turing machine: a mechanism of discrete symbols for algorithmic processing. This remarkable observation places the ontology of digital data in a broader context: discrete forms of data processing predated electronic computers, and these computers historically built on the techniques and even the staff of paper-based data systems (indeed, as Krajewski notes, the library card leads directly to the punch card, and thence to electronic storage) (8). This means that the infrastructure of digitality — the idea that information needs to be broken down into bits and indexed for access — has been evolving for centuries. Digital data is thus an essential foundation of modern knowledge systems, whether analogue or electronic. But this foundation often remains hidden: just as today's software interfaces hide bits and files, the ordered drawers of card catalogues hide a revolution in thinking: knowledge as a set of discrete, manageable records. In both cases, it is only through such records (cards or bits) that larger knowledge structures can be built and navigated.

The chapter attempts to look beyond German media theory to a comparative analysis of digital data and language, exploring their structural, epistemic and interpretative relationships. An important insight of Lev Manovich is that the database has become a new symbolic form in digital culture, inherently opposed to traditional narrative (8). The question is therefore how the

logic of data differs from (and converges with) language. This sets the stage for rethinking hermeneutics in the age of data: if interpretation in the humanities has long been about the interpretive reading of texts, what happens when our "texts" are databases, algorithms and visualizations? We maintain and deepen the debate on hermeneutics by examining how cultural techniques of data are reshaping interpretive practices. In particular, we consider whether the "post-hermeneutic" stance of cultural techniques, which originally sought to free media theory from the burden of purely interpretive (meaning-centered) approaches, can be reconciled with new forms of meaning-making in data-rich environments (Siegert 6).

Finally, I will return to the early work of Bernard Stiegler, which I have already briefly reviewed in the context of the debate between determinism and constructivism. Here, in particular, it will be important to note that the relationship between man and data is grounded in the fundamental philosophy of technology. Stiegler's notions of techné (craft) and episteme (knowledge), as well as his theory of epiphylogenesis (the externalization of memory in technical media), provide a context for digital data. If technology is what makes us human, and our consciousness is made up of externalized memory traces, then digital data can be seen as the latest stage in this exteriorization — a new tertiary memory that both expands and transforms human meaning and experience. Stiegler's early philosophy of the absence of inherent human instincts – Epimetheus' mistake — as an opening for technology to fill the void allows us to ask how digital data as technical prosthesis redefines the horizon of human knowledge and meaning.

4.1 The Dimensions of Digital Data Operations

Digital data operates through essentially discrete, technical processes that make cultural decisions about what matters and how. The cultural-technical perspective starts with the recognition that before "data" becomes meaningful information or knowledge, it is the result of specific operations: encoding, measuring, sorting, computing, storing, transmitting. Kittler's well-known definition of discourse networks highlights data processing as an elementary cultural activity embedded in material means and organizational systems. In the digital age, the triad of selection, storage, processing is largely carried out by computer infrastructures — from sensors that select phenomena for recording (converting stimuli into signals) to databases that store vast stores of records, to algorithms that process records into outputs

What distinguishes digital operations is their extreme scale, speed and formalization. Bernhard Siegert's work emphasizes that cultural techniques work by translating mundane phenomena into a controlled series of discrete operations — in effect, they replace intrinsic differences in reality with chains of operations that validate those differences (Siegert). In digital media, everything that appears continuous or analogue must be broken down into binary distinctions (yes/no, 0/1) and manipulated by logical procedures. Siegert notes that modern cultural technologies often represent a transition from continuity to code: "reality is digitized", gaps and continuities are transformed into "discrete, countable sets of elements" (*Cultural Techniques* 27). Digitization itself is thus an operation of flattening and abstracting reality. As Sybille Krämer notes, digitality predates computers: it lies in the very act of symbolization, which reduces the world into signs (Flattening, 11) For example, the alphabet "is a prototype of a digital system" (11), which breaks down the fluid spectrum of speech into a finite set of characters, just as today's analogue-to-digital converters sample continuous signals into binary data. By conceiving of digitization as a cultural technique of flattening, Krämer resolves the supposed novelty of computer data into a broader historical practice of making the world writable and computable on a two-dimensional surface (11). Counting systems, notation graphs, and tables have always operationalized reality, paving the way for the more comprehensive discretization's of the computer.

Inside the computer, these operations acquire a logic of their own that is often opaque to human observers. Kittler's provocative claim that essentially says what literary studies do is data processing sets aside interpretation in favor of an examination of the machine manipulation

of symbols. He showed how digital media makes all content (texts, images, sounds) comparable as a stream of binary numbers. Any human-readable interface — a text on a screen, an image, a sound — is, as Kittler puts it, merely a "surface effect" of the underlying computations (Kramer on Kittler, 11). Inside the machine, there are no images or words, only tensions and binary states dynamically rearranging themselves. The media that once captured the richness of reality (the detailed image of a photograph or the analog sound of a phonograph) are subsumed under a single code, where "the file on the computer is not inherently sound or image, but can only be interpreted as such on the surface" ("Flattening" 3). This unity of operation — the reduction of all media to computable data — is a feature of the cultural technique of digitization.

It provides unprecedented flexibility (any content can be transformed, copied, remixed or transmitted in the same basic format) at the cost of content that is initially unreadable to the human senses until an interface reverses it. The operational dimension of data thus creates a gap between what the machine "reads" and what we perceive: an infinite stream of bits circulates in networks "without any assumption about the human", as Kittler, following Claude Shannon, notes (Kittler, history, 1). Think of metadata, protocols, handshakes between servers. These automated exchanges are cultural techniques in action: processes that frame human culture while often bypassing direct human interpretation.

4.2 Infrastructures of Epistemology: from Recording to Knowledge

The operational basis of digital data underpins a new epistemology of data — new ways of producing, validating and organizing knowledge. The cultural techniques perspective teaches us to look at the infrastructures, the very conditions of knowledge rather than knowledge: the files, lists, indexes and databases that silently condition what can be known or said. In her book *Files*, Cornelia Vismann outlines the genealogy of filing systems: how legal and bureaucratic knowledge in the West has been built on centuries-old record-keeping techniques, from the Roman acta to medieval registers to modern dossiers. Vismann argues convincingly that "the subject, the state and the law are revealed as the effects of specific registration and filing practices" (Files, xii). In other words, the very notions of individual identity (the subject), political order (the state) and normative rules (the law) were themselves epistemic by-products of an archival technique — the ability to enter and retrieve written records. When these files are converted into digital form (in the form of electronic records, databases and computer files), the same logic applies: the files transmit and process entire systems of knowledge and power (xii). Today we see this in the way databases underpin everything from scientific research to government; knowledge is less derived from solitary minds than from interactions with data warehouses and their organizing algorithms.

Markus Krajewski also highlights the epistemological role of small, technical arrangements in his study of cardboard and databases. By tracing the development of the card catalogue "paper machine", Krajewski shows that the ability to know (for example, the study of literature or the management of imperial government) is increasingly based on the treatment of information as discrete units (8). The file-card system taught users to atomize knowledge into small facts and then recombine them; this anticipates today's data practices, which break information into database fields or data points that can be sorted and retrieved. Epistemologically, this represents a move away from continuous narratives or holistic knowledge towards modular, recombinable chunks of information. As Lev Manovich observes, in computer culture, the database (a structured collection of discrete elements) has become a new cultural form that rivals narrative coherence ("Database" 1). He continues: "The database represents the world as a list of elements... it refuses to order this list", whereas narrative creates a causal trajectory through events ("Database" 8). In the age of data, knowledge often takes the form of a list or set — for example, search engine results, statistical datasets, catalogues —

rather than a story. This list theory of knowledge allows for flexibility and reordering but may lack the explanatory closure of narrative. It is telling that Big Data analytics sometimes boast that it is theory-free: the patterns in the data speak for themselves (or so they claim), without the need for a guiding hypothesis or story. This attitude resonates with Kittler's anti-hermeneutic position: if culture can be understood as data processing, then perhaps meaning (and through its narrative explanation) is an "illusion" imposed on an essentially technical process (*Discourse networks* 369).

However, the appearance of data-based knowledge neutrality — the idea that the database is a transparent reflection of reality — is itself a cultural construct that critical theorists' question. The cultural techniques of data include not only recording, but also classification and modelling. In any given data set, someone (or some algorithm) has decided what categories to use, what counts as a data point, what filtering criteria to apply — all of which shape the knowledge that can be extracted. Siegert points out that cultural techniques work through the creation of distinctions, for example between signal and noise or between one category and another (*Cultural Techniques* 30). In data epistemology, these distinctions are formalized in the form of schemas, ontologies or algorithms that determine what is informative. For example, a social media platform's database may categorize users according to predefined fields (age, gender, location), thus creating a particular, quantifiable vision of the "social." The knowledge generated from such data (say, analyzing user behavior) inevitably reflects these operational categories. The medium of data is never a passive transmitter of truth; it actively shapes truth by the form in which it captures reality. In this sense, empirical knowledge is turned on its head: instead of describing reality, data construct reality.

Kittler's analysis of modern knowledge practices reinforces this point by showing how the boundaries of what can be known are widened or narrowed as the media change. In his view, the universal system of alphabetic writing and printing in the 19th century (what he calls the Discourse Network 1800) created the illusion of author and coherent internal meaning, but the emergence of technical media around 1900 disrupted this unity, with knowledge dispersing into separate sensory channels (sound, image, separately stored text) (117-118). Now, with the digital convergence of the 2000s, we are faced with a new, unified discourse network, where all media streams converge in a single numerical code. Kittler has predicted that this digital convergence will create new forms of knowledge similar to the pre-print era — perhaps a return to a kind of knowledge that is transcendent of the individual and not at all focused on human

authors. Indeed, we are now talking about networked knowledge, collective intelligence and artificial intelligence-generated insights, all of which put the individual thinker at the center. The epistemology of digital data thus tends towards a networked, processual understanding of truth: knowledge is that which is generated from data that is continuously processed by technical systems, often too large for any one person to fully comprehend.

4.3 Digital Data and Language: Structural and Hermeneutical Differences

Is digital data a new language or something fundamentally different? This question guides the comparative analysis of data and language. In this section of my dissertation, we can finally take stock of this question that has been hanging over our heads since the beginning. Structurally, both data and natural language consist of symbols that refer (directly or indirectly) to the world. But their logic is crucially different. Lev Manovich's analysis of databases and narratives summarizes one of the structural differences: traditional language-based narratives are linear and syntagmatic, progressing through a sequence where the meaning of each element is formed in context, whereas digital data (especially in the form of a database) is modular and paradigmatic, a collection of elements whose sequence is not fixed ("Database" 8). He famously noted that "database and narrative are natural enemies", one an unordered list, the other an ordered story ("Database" 8). Narrative makes sense through the interpretive linking of events into a coherent story (cause and effect, theme and variation). Data, by contrast, often makes sense by being collected and organized into non-linear structures — tables, graphs, and networks. For example, a novel (language) and a dataset can both represent the same social phenomenon, but a novel does so through a story, characters and plot, whereas a dataset does so through a matrix of variables and values. A dataset "refuses to order" events as a story would, leaving it to the analyst or algorithm to apply queries or visualizations that produce patterns ("Database" 11).

Epistemologically, language and data also suggest different ways of knowing. Language (in the humanities tradition) is linked to hermeneutics — the art of interpretation, the search for hidden meanings, ambiguities and contextual nuances in texts. The richness of language lies in its polysemy and its dependence on context and connotation. Digital data strive for the opposite: clear, formalized representation. A data point ideally has a single defined meaning (e.g. a temperature value in degrees Celsius or the age of a user in years) to be processed smoothly. In short, natural language tolerates and even requires ambiguity and interpretation, whereas data requires clarity and explicit structure. Sybille Krämer characterizes this as the difference between semantic depth and surface: classical interpretation often seeks what is "unsaid and hidden behind the visible surface", whereas digital methods and notations favor the "sensory signature", that which is explicitly there and which can be directly read or measured. In digital literacies, for example, scholars have argued for "surface reading" as opposed to symptomatic

deep reading ("Flattening" 4), focusing on patterns (word frequencies, co-occurrences) that can be detected on the surface of the text and quantified. This reflects the impact of data thinking on interpretive practice: treating the text as data, flattening it into a two-dimensional array of features, and analyzing these features statistically, rather than looking for hidden symbolism or authorial intent.

But despite these differences, digital data and language also intersect and intertwine. Computer code is often described as a language (syntax, grammar, semantics) — programming languages do indeed mediate between human intentions and machine operations. Kittler went so far as to say that in the age of digital computation, "no human being writes any more... today human writing runs through inscriptions burned into silicon by electronic lithography" ("There is no software" 220) According to him, machine code is a continuation of human writing by other means, with the crucial difference that its primary reader is the machine, not the human. Code and algorithms can be seen as a grammar of data that structures the way data is generated and processed. Meanwhile, human language is data-driven: natural language processing turns words into number vectors, sentences into parseable tree structures, semantics into computable ontologies. The boundary between language and data is thus blurred in fields such as computational linguistics and artificial intelligence, where language itself becomes a data set that machines learn. Manovich's idea of transcoding captures this mutual infiltration: cultural content (such as images or stories) is translated into the computer's native language, numbers, and vice versa, the outputs of data are translated back into the cultural forms (visualizations, narratives) we know ("Database" 11). Digital data for Manovich is thus not a negation of language, but a metamorphosis of it within technical constraints — a kind of meta-language optimized for predictability and interoperability at the global level.

One productive way of comparing data and language is to examine interpretation. Linguistic interpretation is traditionally open: a poem or a novel can be read in many different ways. Interpretation of data, on the other hand, often tends to be closed: a single analysis (e.g. a statistical result or a graph) that "speaks for the data." However, as data scientists point out, this is an illusion — data also needs to be interpreted, and can be sliced in different ways. The difference is that the act of interpreting data usually involves an intermediate formal step: choosing which algorithm or query to run, which visualization to draw, and which model to use. This makes the interpretative act explicitly procedural. In fact, interpretation itself becomes a kind of cultural technique when dealing with data. Johanna Drucker argued that in the

humanities we should treat data as capta (that which is recorded) to emphasize that data are constructed and need to be interpreted, not simply given as truth ("Humanities approaches", no page.) The structural form of the data (rows and columns, vectors of features) requires that meaning is extracted through formal operations — clustering, correlation, mapping — while language requires narrative and exegesis. ("Humanities approaches", no page.)

As a result, we some tendencies what we might call, with a little bit of exaggeration algorithmic hermeneutics: reading patterns in data as if we were reading meanings in a text. Remote reading of literature, such as stylometry for example, counts the frequency of words in thousands of novels and interprets trends; it is a combination of data analysis and literary interpretation that bridges the gap between list and story by turning lists into stories (e.g. a narrative about how the vocabulary of a genre changes over time).

I would argue that Lev Manovich's insight that the database has become a cultural form does not mean that narrative is dead; rather, it means that narrative is emerging in a different way, often post hoc. ("Database" 1). In cultural analytics or journalism, we see data being turned into narratives — data is collected and analyzed, and then a story is written to explain the results. In traditional humanities, in a way, the narrative was the data (the primary source of meaning); in data analytics, the narrative is a secondary interpretation of the prior data structure. I argue the trend was probably most evident during COVID-19 when case numbers were globally published almost real-time, while experts, commentators and media outlets were looking for the possible causes for the change in the numbers. This reversal underscores the epistemic shift we are undergoing: the ambiguity of language is complemented by the predictability of data, and conversely, the silence of data (the absence of inherent meaning) is complemented by the contextualizing and explanatory capacity of language.

Comparing digital data and language also reveals their common ground: both are systems of discrete symbols that require a reader or interpreter. Sybille Krämer's work on operational writing suggests that historically, notes (from musical scores to mathematical formulae) have served as an actionable language, instructing performances or calculations. In the digital space, data can be seen as an operational signal — they become meaningful when acted upon by software (just as a musical score comes to life when played) ("Writing, Notation" 522). The structural rigor of data (like the strict tempo of a score) allows certain operations (calculations, visualizations) while restricting others. A more free-form language allows for creativity and ambiguity, but without formalization it is harder for machines to use.

Maybe a bit of an oversimplification, but for now we can say: language is the medium of cultural exchange between humans and humans, while data is the medium of exchange between humans and machines. Data is "written" by machines or humans in a form that machines can read; machines then "read" (process) it and produce results that humans can read. This three-way loop (human writes → machine reads → machine writes → human reads) is the new communication circuit of our time, distinct from direct human dialogue mediated by natural language.

The rise of data-driven methods has led to a rethinking of hermeneutics, the theory and practice of interpretation. Classical hermeneutics (from Schleiermacher to Gadamer) was developed for texts, assuming that meaning could be discovered through careful, context-aware reading. However, when faced with large data sets or algorithmic outputs, the traditional hermeneutic circle — the movement between parts and whole in order to grasp meaning — seems to break down. How do we interpret the pattern that an algorithm finds across millions of data points? How do we deal with reports that are not directly written by a human author but generated by a machine process? These questions have led to calls for a new "digital hermeneutics" or hermeneutics of algorithmic results. In a way, the cultural-technical approach offers a provocative position: it suggests that what we call interpretation is itself behind technical practices. For example, reading a book is not merely mental, but relies on the physical techniques of printing, the alphabetic writing system, and even the act of turning the page. The shift to digital media can therefore transform interpretation by transforming the underlying techniques of reading and analysis.

One notable development is the idea of post-hermeneutic analysis in media theory. Bernhard Siegert notes that when German media theory entered its "second phase" in the late 1990s, there was a "conceptual transformation of media into cultural techniques" that could be called post-hermeneutic ("Grids, Filters" 6). This shift was aimed at going beyond the simple decipherment of media content (the hermeneutics of message and representation) to examine the processes and materialities that make these messages possible. In doing so, he "freed media and technology from the burden of having to play the bogeyman of hermeneutics" by framing media technology not as a threat to meaning but as a productive of meaning ("Grids, Filters" 6). From a practical perspective, a post-hermeneutic approach would look at how an algorithm orders search results (the cultural technique of ranking) rather than what a search result means. Flattening is understood as the way in which computer analysis projects multidimensional data

(including texts) onto two-dimensional surfaces such as screens, tables or graphs . Interpretation thus does not dive deep into the text but literally begins with the examination of surfaces — patterns, distributions, visualizations. This methodological shift values pattern recognition over deep reading. The "unspoken and hidden" depths traditionally sought by hermeneutics are eclipsed in favor of observable, countable features ("Flattening" 3).

However, the tension between data analysis and hermeneutics can be productive rather than antagonistic. Humanities scholars increasingly adopt a hybrid approach: they use algorithmic tools to obtain a remote view and then apply interpretive insight to understand why these patterns might be important. For example, a topic modelling algorithm may discover clusters of words in a corpus of novels (a non-intuitive pattern that the computer "reads" for us); the scholar must interpret what topics or discourses these clusters represent. Here, the cultural technique (topic modelling) transforms hermeneutics by limiting it — one interprets the output of the model, not the raw text. At the same time, it extends hermeneutics by opening up new "wholes" (complete archives) that were previously beyond the human ability to read. In effect, the hermeneutic circle is extended to include machine processes: part-part interaction can now exist between individual data points and a computed model of the whole data set. Meaning is generated in the interaction of human curiosity and machine sorting.

Scholars of cultural techniques like Vismann and Krämer, with their anti-foundationalist stance, warn us against re-importing naïve hermeneutics into the digital context. Vismann's work is itself "anti-hermeneutic," as it focuses on the material media of law (files, dossiers) rather than on the interpretation of legal texts or intentions (Minkinnen 1606) This provides an important lesson: sometimes the absence of interpretation is itself a technique. Deletion, forgetting, ignoring — these are also operations that shape meaning by negation. Digital systems have their own built-in hermeneutics, in the sense that algorithms "decide" what something means to the system: for example, a spam filter interprets certain email content as undesirable based on keywords. However, these are operational hermeneutics, often implicit and inaccessible to users. One could say that hermeneutics in computer systems are partly automated. The danger is that this automation is opaque — a problem often raised in the ethics of artificial intelligence (the "black box" problem). For the humanities, this opacity requires a kind of metahermeneutics: interpreting the interpreters (the algorithms) and exposing their biases and assumptions. The cultural techniques approach, which focuses on the basics, is well suited to this task. It invites us to interpret the technical terms themselves.

Thus, hermeneutics in the age of data becomes a double project: interpretation through data and interpretation of data systems. The first is when scholars use data as a means of interpreting culture (e.g. mining social media data to understand the public discourse at). The second is when scholars interpret how data systems themselves function culturally (e.g. analyzing how Google's search algorithm shapes knowledge). Both are critical in contemporary humanities studies. The first extends the range of data that can be interpreted (vast corpora, multimedia archives), while the second makes interpretation reflexive, ensuring that we understand the frameworks we use. In this light, data cultural techniques do not end hermeneutics but reconfigure it. They force hermeneutics to reckon with new actors (machines, algorithms) and new texts (data sets, visualizations).

To illustrate this reconfiguration: think of the notion of the hermeneutic circle — the idea of understanding the text as a whole by reference to its parts, and vice versa. In the context of big data, the "whole" might be a database, and the "parts" might be individual entries. A researcher can formulate a hypothesis about the whole (say, a trend in the data), test it by analyzing the parts (perhaps by sampling or looking at outliers), and revise the hypothesis — in a way analogous to traditional hermeneutics, but now using computational techniques such as sampling or visualization. The circle can become a tool-mediated spiral: initial interpretation informs data retrieval, which yields results that provide information for further interpretation. This process can be observed in iterative data analysis in, for example, digital history or sociology. Hermeneutics is thus preserved in addition to data operations.

In fact, a radical solution can resolve the dilemma: hermeneutics itself can be seen as a cultural technique — a learned practice with tools (from the concordances and dictionaries of classical philology to today's statistical software). The humanities are adapting by creating new technical practices of interpretation: algorithmic criticism, remote reading, network analysis of texts, etc. All of these are a kind of fusion of the hermeneutic impulse and the operational rigor of data processing. Importantly, these developments underline that interpretation has not become irrelevant, but rather more collaborative — between humanists and machines, theory and process.

We now consider Bernard Stiegler's perspective on this situation. If data hermeneutics represents a new chapter in the co-evolution of humans and technical memory systems, then Stiegler's philosophy can provide a longue durée context. His focus on technology as a

conditioning milieu for human experience will help to articulate why and how digital data matter not only functionally, but existentially and historically for humanity.

4.4 Techne, Episteme and Epiphylogenesis

At this point, we need to return to Stiegler's idea of the prostheticity of the human being in order to frame the image of the human being created by the data. This approach can frame our understanding of the power of digital data to shape culture and knowledge, and ultimately to shape the human subject. As reviewed in Section 2.3, in Technics and Time 1: The Fault of Epimetheus, Stiegler takes as his starting point the myth of Prometheus: Epimetheus, who distributed abilities among all creatures, forgets to give humans substance or natural advantage, leaving them naked and vulnerable. Prometheus compensates by endowing humans with technology (fire, knowledge). Stiegler interprets this to mean that man is essentially a technical being, defined by an inherent deficiency (lack of instinct, Instinktarmut) which he must make up for by artificial means. From the very beginning, therefore, man externalizes his being into tools and symbols. He calls this process epiphylogenesis: the evolution of life accumulated in technical vehicles outside biological organisms. Put simply, epiphylogenesis is the idea that human cultural memory and knowledge is stored externally: in devices, writing, eventually images, now digitally, and that this external memory evolves over time, reflecting back on human development (*Epimetheus* 140-141).

Using Stiegler's framework, digital data is the latest form of epiphylogenetic memory. Briefly and concisely, focusing on the parallels, we can say that while the invention of writing enabled the capture and transmission of knowledge beyond the lifetime of the individual, and printing enabled the explosion of collective knowledge, digital storage and computation now exponentially expands the scope and speed of memory transmission. It is important to note, however, that this is not a self-evident and by no means neutral process. Stiegler's important insight is that technology is not merely an adjunct to human life but constitutes human temporality and knowledge. It is impossible to separate the human from the technological, he argues; humans are defined by our inherent technicity, which emerges at the same time as we become human (Bluemink, no pag.) In this light, the vast data infrastructures of our time (archives, cloud servers, artificial intelligence models) are not alien to humanity, but rather an integral part of our episteme, our way of knowing. To use Stiegler's term, they are tertiary preservation: memory objects outside of us that carry collective experience (*Epimetheus* 140) A simple example is how Google (as a collective memory bank) has changed the way we recall facts — we often do not memorize information if we know it is stored online. At a deeper level,

"the newborn child is born into a world in which tertiary preservation precedes and awaits him" (Siegert, qtd in. Bluemink, no pag.). The child is born into a world saturated with digital data (photographs, medical records, social media, sensor data) that form the backdrop of his social reality. In other words, digital data are part of the horizon of meaning in which the subject can be created.

Stiegler's distinction between techné (skill) and episteme (knowledge) is also blurred in the case of digital data. Traditional philosophy often considered techné (practical knowledge, making) inferior to episteme (real knowledge, science). Stiegler, however, following Heidegger and Simondon, sees techné as a condition of episteme: we have stable knowledge only because we exteriorize and repeat practices by technical means. In the context of the process of datafication, this is manifested in the fact that scientific knowledge now depends on data-collecting instruments and analytical software (technics) — there is no large-scale knowledge without large-scale data technics. One could say that data is the place where the techné and the episteme converge. Creating a database (techné) is also an exercise in organizing knowledge (episteme). For example, the design of an algorithmic model requires technical skill, but also encodes theoretical assumptions about the domain (knowledge). Stiegler's work suggests that rather than treating the technical and the epistemic separately, we should consider them as a continuum: how knowledge (savoir-faire) and knowing (savoir) co-evolve (Bluemink, no pag.). Digital data practices exemplify this co-evolution. The creation of huge searchable image databases, for example, has changed art history (artworks can be compared to scale) — but the creation of these databases has required technical mastery and in return has created new art historical knowledge.

The notion of organized inorganic matter, which appears in Stiegler's narrative, is particularly well suited to data. He describes technology as an "organized inorganic" support on which human beings rely — like a skeleton outside the body (Tinnell 134). Data centers and computer chips are literally organized inorganic matter (silicon, metal, plastic) configured to store and process symbols. These machines now embody a significant part of human memory and decision-making. As we have seen in *The Fault of Epimetheus*, Stiegler often returns to the example of writing as the "exteriorization of memory" — and what is a database if not writing by other means? But the scale and automaticity of digital data introduce a new dynamic: real-time epiphylogenesis. Information flows globally in real time, and algorithmic feedback loops can change human behavior almost instantaneously (think of how social media trends influence

real-world actions within hours). Stiegler's concern in his later works is that this acceleration could lead to the loss of knowledge about life (savoir-vivre) and knowledge about action (savoir-faire), as more and more is delegated to machines (Bluemink, no pag). However, the emphasis in early Stiegler is on the fundamental point: that this delegation is inherent. It has always been the case that the adoption of a new technique changes us. The challenge is to appropriate digital technology in a way that supports human individuation (the growth of knowledge and spirit) rather than short-circuiting it.

Stiegler's ideas on temporal objects and tertiary conservation also provide insights into interpretation. A book or a film is a temporal object that can be replayed, reread, thus allowing the past to be "preserved" and interpreted later. Digital data, because it is durable and easily copied, extends this phenomenon — but also transforms it. The sheer volume of data retained means that we cannot individually experience all of our tertiary preservation (who can read all of Wikipedia or even all of their email?). We therefore rely on techniques (search algorithms, recommender systems) to mediate our access to tertiary memory. Stiegler might also say that our choice of what to pay attention to is increasingly left to technology, which affects how memory and attention work (a topic he later explores in terms of the attention economy). But early Stiegler would put it more neutrally: this is a new stage in the grammatization of knowledge. According to him, grammatization is the process of breaking down the flow into discrete units, such as writing did with speech (Tinnell 134). I would argue that digital data is grammatizing more and more of life — turning it into discrete bits — and thus creating a new layer of epiphylogenetic memory that both carries forward the past (recording everything) and shapes the future (through predictive analytics etc.).

Stiegler basically prompts us to ask: what does it mean for humanity to entrust so much of our collective memory and abilities to digital data systems? His answer would revolve around pharmacology: technology as both poison and cure (Bluemink, no pag.). Focusing on Stiegler, I would say: it means that the human adventure has always been one of interaction with external memory carriers, and digital data is just the latest milieu in which the human adventure continues. The task, then, is not to reject or fear data, but to understand it — to continue a digital epoch (suspension and reflection) on how these data technologies are reshaping our existence and our knowledge. Stiegler's emphasis on technology as a horizon of desire and knowledge suggests that we need to cultivate a care for our digital techniques if we are to use them for positive individuation. Already in his early work, he warns that forgetting the techné would

lead philosophy astray (Bluemink, no pag.). More broadly, forgetting the techné behind digital data — treating data as a given rather than a created thing — would lead our contemporary thinking astray. Instead, we should acknowledge that thinking in the 21st century means thinking with, through and about digital data as a condition of who we are.

4.5 Conclusion

If we make digital data the object of theoretical study, we find that it is much more than a technical by-product; it is at once a cultural operator, a cognitive engine and a symbolic form. Through the lens of cultural techniques, we have seen that digital data embody sequences of operations that render the world in discrete concepts, extending the history of symbolic practices from the first alphabets to today's algorithms. I have shown how these operations — selection, encoding, storage, processing — underpin new ways of knowing, as databases and networks become the scaffolding of the contemporary episteme. By comparing digital data and language, we have highlighted the tensions between structure and meaning: data, with its list-like logic and machine-readability, challenges the linear, meaning-laden domains of language, even as new hybrid practices unite the two. The hermeneutic traditions of the humanities are not abolished but transformed by data-driven approaches, shifting the focus towards surface patterns and procedural interpretations without, one hopes, abandoning the search for meaning altogether.

The chapter underlined that interpretation itself becomes a cultural technique adapted to the age of algorithms. As I tried to show, an important element of Stiegler's philosophy of technology is the idea of the technicization of memory itself, which he describes in terms the of epiphylogenesis (*Epimetheus* 140). According to this concept, human knowledge and memory are not only accumulated through biological (genetic or individual neural) pathways but are also stored and transmitted in external technical media — from cave paintings to books to digital data. The argument is then this tertiary memory that has enabled the accumulation of humanity's collective knowledge over generations: every technical innovation is also a materialized memory of a slice of human experience. Thus, human history can be understood as a history of technical forms of memory (drawing, writing, printing, digital storage).

Ultimately, the cultural techniques of digital data emerge as constitutive elements of our contemporary condition: they are the often-invisible practices that enable data quantification, the epistemic scaffolding that supports new knowledge systems, and the symbols through which much of the meaning now flows. By making visible and theorizing these techniques, we prepare ourselves — as scientists and citizens — to engage more thoughtfully with the digital data space. I would argue it is important to learn to see digital data processes not as an alien code, but as a different kind of cultural technique. Interacting and understanding these cultural processes requires the development of new approaches. Just like a text that requires new

analytical tools and renewed hermeneutic care to read. In doing so, we take forward the project exemplified by our German theorists and Stiegler: the relentless examination of how our tools and media shape who we are and what we know.

5. Case Studies: The Cultural Techniques of Digitality

The last, but also longest chapter of my dissertation aims to demonstrate through specific case studies how my theoretical framework and broader conceptual considerations outlined in the previous chapters can be applied to specific cultural phenomena. In this regard, I would like to achieve two goals. First, I aim to use the interpretive framework I have developed to make statements about digital phenomena that demonstrate not only the applicability of the framework but also its critical productivity. Second, I would like to bring the underlying interface of digital cultural phenomena into the horizon of interpretation of the humanities. In doing so, my intention is to open the floor for humanities to interpret and critically reflect on today's digital phenomena, especially in the context of Hungarian universities, as most programs focus overwhelmingly on classic texts and traditional philological approaches.

The central thesis of the theoretical section of my dissertation is that the concept of digitality, as an epistemic approach, can and should be applied to the subject of humanities' analysis. In other words, it would be reductionist and potentially misleading to consider a video game, for example, solely as entertainment software or an independent aesthetic creation. Rather, it should be interpreted as a culturally and technologically embedded medium. A proper interpretation can only be revealed by taking into account the medial conditions specific ontologies of digitality, as, which is embedded in digitality from the outset.

The three case studies presented in this chapter share the common feature that they interrogate cultural phenomena that are inherently digital, can only come into being within digitality, and have also developed within it. This approach allows me to show that the digital is not simply an external environment, but a constituent element of culture through its cultural techniques.

The perspective of cultural technologies is therefore equally important. It is essential to ask the specific digital data operations are being performed in these case studies and how these operations are transformed into processes that become understandable and interpretable to humans and ultimately create cultural forms. This set of questions emphasizes that digital culture is not only about content, but also about the processes, transformations, and infrastructural logics that underlie and enable cultural production.

It is also important to note that the case studies I present here always refer to specific cultural objects, rather than to the medium in general. There are methodological reasons for this

choice. The conceptual horizon outlined in the previous chapters, primarily the data-based ontology of digitality and the role of cultural techniques, can be considered generally applicable, but these connections become productive in the practical sense when examined through specific examples. Beyond the theoretical framework outlined in the previous half of my dissertation my intention is not to establish general trends in relation to specific mediums of digital culture, but rather to show, through the analysis of specific objects, how the cultural techniques of digitality can be detected and how a given phenomenon reveals the logic of datafication.

The case study as a methodology offers a suitable framework for this. In his classic work, Robert K. Yin emphasizes that case studies are not merely illustrations, but one of the most effective tools for examining real-life context-dependent processes, and also when boundaries between phenomena and context are not clearly evident (13). This is particularly important in the study of digital culture, where phenomena always arise at the intersection of technical-infrastructural and cultural contexts. The case studies that I am presenting in this chapter thus not only "apply" theory but also reveal the micro-level operations that reveal the macro-level logic of digital culture as a whole. The shift from theory to practice can therefore be interpreted as a two-way process. On the one hand, the theoretical framework provides the conceptual and interpretative perspectives on which I approach the case studies. On the other hand, the specific analyses feed back into the theory: they reveal its strengths and limitations and create opportunities for critical refinement.

In the first case study I read League of Legends champion skins as born-digital commodities with their very being defined by data operations. In a datafied ontology, skin is not just a cosmetic feature. It is a bundle of describable attributes, version histories, and recomposable assets that remain addressable across different media. That is the reason the same skin can anchor a character arc in Arcane, trigger mechanics in an in-game event, or reappear in a music video without losing identity. I further argue that it also acts as a pharmakon: as "medicine," skins externalize memory and deepen identification; as "poison," the same addressability feeds metrics, surveillance, and scarcity loops that optimize attention. Aesthetic coherence, then, is not only art direction but schema discipline: the brand stays consistent because the data model makes every variant interoperable by design.

In the second case study, concerned with the project of the Hungarian Database of English Studies (HADES), I attempt to reconcile traditional philology with computation by taking "data hermeneutics" seriously. Bibliographic records become triples and graphs (RDF,

SPARQL), but I treat each modeling move—class choice, relation type, disambiguation—as an interpretive claim about the field's ontology. In other words, the database is not only a repository of information: it constructs categories and puts scholars, texts, and translations into machine-readable relations that can be queried and recombined. The rediscovery of Kosztolányi's lost Byron translation shows how a datafied ontology makes latent links retrievable at scale: once the entities and relations exist as addressable data, new constellations appear. The HADES project attempts to shifts infrastructure from passive archive to cultural agent, where the shape of the graph actively guides what counts as philological knowledge.

Finally, I compare one famous example of print culture, Shakespeare's First Folio to a famous example of AI platforms, ChatGPT to show how media remake the ontology of knowledge creation and our relationship to texts. Print stabilized Shakespearean texts by assigning durable identifiers: original texts, formal categories, signatures, so works could be cited, compared, and archived; the text's ontology became fixed objects. LLMs, by contrast, inhabit a datafied ontology of parameters and token probabilities: every output is a contingent state computed from a model snapshot, not a stable edition. That probabilistic, versionable-yet-evanescent status reintroduces philological instability reminiscent of pre-print orality, unsettling authorship and originality. The question then is how our relationship to concepts such as originality, authorship and established knowledge changes once again. Also, I will attempt to outline an inquiry into the concept of truth and how AI-based platforms problematize public discourse.

5.1 Born-Digital Global Commodities as Transmedia (World)-Building Blocks: League of Legends' Champion Skins

To present an interpretation of the born-digital League of Legends transmedia universe, I treat digitality not merely as a technical backdrop but as an ontological condition. I argue that champion skins are digital objects whose mode of existence is determined by the formal operations of describability, versionability, and recomposability. This datafied ontology enables the same object to "leave a trace" differently across media surface: as an audio-visual apparatus in the game client, as a narrative node in videos and comics, and as an identity marker in esports events.

This mode of existence presupposes a series of digital data operations: capture, labeling, and indexing (skin lines, champion categories, event-tied releases); aggregation and recombination (bundles, variants); version control and feedback (continuous patches, visual and sound-effect updates); and the differentiated regulation of access (in-game economy, event participation, entitlements). A skin is thus not merely an aesthetic surface but also a data structure: an entity describable through attributes and metadata, measurable and algorithmically distributable, serving as a unit within both platform operation and story construction.

Concurrently, on the user side, cultural techniques are activated that organize practices around the digital object: collecting and curatorial arrangement (collections, favorites); performative display (streaming, highlight videos, social posts); and identity work and communal affiliation (clan and series attachments, the aesthetic codes of events). These techniques describe not only patterns of consumption but also modes of meaning-making: the skin functions as a building block of the narrative world, a modulator of gameplay experience, and a coupling point for economic and circulatory logics.

The digital cultural trend of the 21st century is becoming increasingly clear: as a result of media convergence, our cultural products have increasingly become transmedia storytelling platforms, whose narratives are presented to the audience through a variety of media, including film, music, text, and comics. Video games are no exception. One of the world's most popular online video games, League of Legends, is a good example of this trend, with an ever-expanding universe that includes animated series, fictional music bands, animated videos, comics, esports events, and other media. My argument is that at the heart of League of Legends' transmedial world-building strategy are so-called "champion skins": these are essentially visual accessories

that change the appearance of playable characters, called champions. However, champion skins are far from trivial accessories; they are digital objects with narrative significance, which makes them important from both a technical and a humanities perspective. This significance lies in the fact that champion skins act as a kind of glue, connecting storylines that span different media and maintaining players' commitment to the ever-expanding narrative of the video game universe. My analysis is based, among other things, on Yuk Hui's ontology of digital objects, Derrida's concept of trace, and Stiegler's theories of technology as pharmakon. In this chapter, I attempt show how skins become channels of story and experience, highlighting the positive aspects of transmedia storytelling, while also reflecting on the problems of commodification, digital surveillance, and exploitation.

League of Legends champion skins play a remarkable dual role: they are both digital objects—created, updated, and distributed using code—and cultural texts—filled with story details, fan engagement, and economic value. This dual nature makes them an ideal focus for exploring how narrative, technology, and capital intersect in the contemporary media world. This chapter examines each dimension of this phenomenon and builds a case that skins are key technocultural objects through which the League of Legends franchise simultaneously tells a story and conducts business. In doing so, I hope to contribute to a more nuanced theoretical perspective on video games as transmedia systems, as well as to the possibilities and tensions that arise at the intersection of art, narrative, and business.

The concept of transmedia storytelling was coined by American media scholar Henry Jenkins, who defined it as referring to narratives that "unfold across multiple media platforms, with each new text making a distinctive and valuable contribution to the whole" (95-96), which became the standard definition of the phenomenon. Since the introduction of the term other researchers contributed to the understanding of transmedia. Most notably Marie-Laurie Ryan argues that transmedia storytelling is not a single story-arc dispersed through various media. Instead, transmedia stories are held together by taking place in the same story world, therefore it would be more accurate to refer to this strategy as transmedia world-building (4-5). In line with Ryan's thoughts, I would also base my conception of League of Legends transmedia world as acts of world building relying on data ontologies.

The fictional universe of League of Legends expands across various media: an Emmy-winning Netflix series (Arcane), chart-topping music collaborations (K/DA, True Damage), comics and short stories, and a developed esports scene. Champion skins have become key

nodes in this transmedia network. The virtual pop band K/DA was introduced in the viral music video "POP/STARS" (2018), in which players could purchase K/DA skins for Ahri, Akali, Evelynn, and Kai'sa, the champions who make up the band. Examples like this show that a champion skin is more than just a visual backdrop. Rather, they are transmedia objects that can carry cross-platform narrative content, aesthetic styles, and cross-promotional potential.

5.1.1 Digital Objects, Data and Ontology

In his article titled "Ontology of Digital Objects," philosopher Yuk Hui argues that digital objects, such as a YouTube video, a Facebook profile, or even an element in a game, cannot be understood as mere inert bits or visual devices — rather, they exist through relationships: relationships of data structures, metadata, networks, and temporal updates (390). The identity of a digital object is derived from its relationships within a larger technical ecosystem and its evolution over time. Using Hui's framework, a *LoL* skin can be considered a paradigmatic digital object. At the technical level, a skin is composed of data files (3D models, textures, animations, sound effects) formalized by metadata schemas (380), such as the skin's identifier in the game database, its name and description, tags associated with the champion and a thematic skin line, or even its price category. These metadata are not separate from the object, but are integral to it, as Hui notes, in digital objects, "ontologies" (data schemas, classification schemes) are closely linked to the data itself (390). A champion skin is essentially defined by its attributes in Riot's content management system, and by how it is indexed in the client's store or the player's inventory.

Perhaps even more importantly, a skin only "exists" for players through its network of connections: it appears on the game client's storefront, promotional media on the video game's website, YouTube spots, wiki and fan pages, and of course within the game's matches, where other players see your champion in that look. It is less a standalone file and more a node in an extensive digital ecology. The ontology of a skin includes not only its code, but also its relationships. Linked primarily to a specific champion — a Lux skin is meaningless without Lux, other skins in the same thematic universe — Star Guardian skins share common visual motifs and storylines, as well as events or narratives that put them into context. For example, "PROJECT: Ashe" is part of the PROJECT cyberpunk themed storyline and appeared during a themed event; it is linked to other PROJECT skins and the storylines of the event through metadata and narrative. According to Hui, a *LoL* skin is a good example of how digital objects are "composed of data and formalized by schemas or ontologies" and how they "permeate our everyday lives on the internet" ("What is a Digital Object?" 380), or in this case the player's everyday life in the game's online community.

One crucial aspect that Hui highlights is the timelessness of digital objects. Because they are created in a dynamic technical environment, they are subject to constant change — updates, version changes, modifications through user interactions. Hui writes that digital

objects should be approached through the temporal relationships created by "the artificial memory of data" ("What is a Digital Object?" 390). League of Legends champion kinetic objects do have a life history. They are often developed and modified iteratively: textures are reworked, new chromes (color variations) are added, models are updated for compatibility when the base champion is revived. Riot Games even occasionally does "art revisions" on old skins to bring them up to modern standards or storylines (for example, when the champion Caitlyn received a model update, all of her skins' art was refreshed to ensure consistency. Skins are also available on a cyclical basis — some are removed from the shop (they go into the Legacy vault) and then brought back later during special sales, giving them a temporary rarity aura. Furthermore, each skin accumulates a kind of metadata archive over time: players' memories and associated notes (screenshots, forum discussions, maybe personal memories: "I got this skin during the 2016 Harrowing event"). In this sense, a skin is not a static object, but a persistent digital entity that persists over time, accumulating layers of history. It contains traces of the moment of its introduction (often commemorating a particular event or media contact) and can be reinterpreted later. It can be said that a *LoL* skin exists ontologically as a concrete set of data (the current version of the game) and as a set of traces of past states and contexts. This is in line with Hui's view that the evolution of a digital object over time, as conditioned by "artificial memories" (logs, updates, stored data), is key to understanding its existence (What is a digital object? 390).

By identifying champion skins as digital objects, I also focus on their programmability and repeatability. Unlike a physical commodity (say, a figure), which once sold is fixed in form, a digital skin can be continuously modified by the developer or even by players (to a limited extent, for example, through user-created mods or custom UI interfaces, although *LoL* does not officially support modding). Riot developers actively refine the code that defines skins, optimizing their performance or enhancing their impact. Each time the game client is improved, skins may be silently modified. In addition, new skins often build on previous ones — reusing rigging or animation code, following popular design patterns — resulting in an iterative design culture where digital assets evolve through the franchise timeline. This mirrors Simondon's notion (conveyed by Hui) that technical objects have an evolutionary trajectory ("What is a Digital Object?", 384-385). One can see the evolutionary trajectory of *LoL* skins, for example, in the way that the earliest legendary skins introduced novel features that later became standard in epic skins, or that the concept of "Prestige Edition" skins (special gold/white versions for collectors) began with K/DA Kai'sa and is now a regular part of Riot's monetization strategy

(Heath and Barovic, no pag.). Each look is thus situated in a network of technical and creative connections: it inherits from previous devices and sets precedents for future ones.

In summary, approaching *LoL* skins from Hui's perspective reveals them as ontologically rich digital objects. They are not mere epiphenomena of the game's code, but integral parts of a relational database that contains the context of stories and experiences. They exist simultaneously as (1) metadata-structured data (code + schema), (2) nodes in a network of platforms and social discourses (game, websites, fan communities), and (3) temporal artefacts with version histories and iterative futures. This understanding sets the stage for our next section: if looks do indeed carry such relational and temporal complexity, then we can assess how they become carriers of narrative traces across media. I now turn to Derrida's concept of the trace in order to theorize champion skins a carrier of transmedia memory.

5.1.2. Trace and Transmedia Storytelling

In Derridean deconstruction, the trace is a sign of absence — an echo of something that is no longer present but still shapes meaning. Derrida described the trace as "a sign of the absence of presence" (Spivak xvii) an elusive reminder that every signifier transcends itself, pointing to previous contexts and deferred significations. This means that if we consider each skin as a sign of a complex sign system (the transmedial narrative of *LoL*), we can assume that the skins carry traces of narratives and media moments outside the immediate game. They are designed artefacts, but also mnemonic vessels: each skin design often alludes to a backstory, event or cultural reference, inviting players to recall or imagine a larger narrative in which a particular visual representation has meaning. In short, skins function as intertextual clues in the transmedia storytelling of *LoL*, connecting different texts (game story, film clips, Arcane episodes, esports events) through visual and thematic references.

To take a concrete example, Jinx's character has a skin called Zombie Slayer Jinx, which depicts the character in a stylized, post-apocalyptic setting and appearance. To an uninformed player, this is just a cool alternative look. But for a fan familiar with *LoL* history and media, this skin might evoke Harrowing (*LoL*'s Halloween event) or be associated with the zombie apocalypse tropes of games and movies. Now let's compare it to Arcane Jinx, a look released in conjunction with the Netflix series Arcane. This skin is essentially the Jinx we see in the series — younger, with distinctive clothing and hairstyle. For players who have seen Arcane, the skin is filled with clues to Jinx's narrative arc from the series: his relationship with Vi, his trauma, and his transformation. The skin is a remnant of the animated series, which now appears in the game. To use Derrida's term, it is the presence of an absence — the story of the series is not literally retold in the game, but the look bears the trace of an absence that still speaks. We cannot see Jinx's emotional journey in the look itself; what we see is an absence (only the visual appearance of it). Yet this appearance triggers the memory of the series' content for those who have the context. The skin thus exemplifies how the transmedial elements of the game function as clues: they refer meaning to another text. The full meaning of Arcane Jinx is deferred to the Arcane narrative — the skin invites us to supplement our understanding by recalling (or viewing) the series. These dynamics follow what Derrida articulates: the trace defies the self-enclosed, fixed meaning, pointing instead beyond itself. Skins, as traces, point to stories told elsewhere beyond the content that can be played directly.

In fact, the looks themselves often serve as a trace, or in other words a 'visible indicators of a missing presence': when a player sees another user in the game using, say, the Azir skin of a 2022 World Cup (assuming there was one celebrating that year's tournament), that skin is a visible indicator of the presence of an entire event (the 2022 World Cup) that is not actually happening in the game. The event is not present in the here and now, yet it is present through the imagery of the skin (perhaps the skin shows the team colors or the motif of the championship trophy). This ability of the skin to contain and disseminate narrative or historical references is fundamental to the transmedia construction of *LoL*.

Riot Games deliberately creates many skins to relate to events in the story or external media. For example, the Pulsefire skins introduce the champions in a time-travel sci-fi scenario linked to the overarching story of a future timeline; the Star Guardian skins place the champions in a "magical girl" anime narrative separate from the main story, with episodic stories linked to it on the Universe website. Each skin, when equipped, carries with it traces of these narrative connections — a player wearing a Star Guardian skin can recall the corresponding short story in which that character fought alongside his team against the darkness. Even without explicit in-game storytelling, the design elements themselves are evocative: Star Guardian skins use shimmering pastel effects and thematic symbols that reference the anime magical girl genre. It is a classic transmedia technique for each medium to do what it does best and visual design is the primary narrative mediator in games. The look becomes a piece of the story world that players can wear, literally embodying the crossover narrative on their avatar.

Transmedia scholar Henry Jenkins notes that ideally, each component of a transmedia franchise should be sufficiently self-contained to be enjoyable on its own but should also form a richer whole (96). The *LoL* skins illustrate this well. The player does not need to know anything about Arcane in order to use and enjoy the look of the Arcane Jinx skin — the skin as a "stand-alone," purely cosmetic item. However, for those who are familiar with the series, the skin provides a richer experience: they can effectively play out the series scenarios or appreciate the subtle design elements (such as the powder burns on his clothes that hint at the explosion in the story). In this way, skins function more as a transmedia extension than as a simple promotional item. They are not merely advertisements for other media (although they do have marketing value) but actually integrate the content of the other medium into the game universe in an interactive form. The result is a kind of distributed storytelling, where narrative

information is spread across platforms, but the in-game look and feel carries enough traces of that information to evoke the whole.

Derrida also talks about the trace being about an absent origin — "reminding" us that the origin was never fully present. In the case of *LoL*, it could be argued that there is no single authoritative narrative account of the lore; instead, the characters' "origin stories" are scattered throughout the biographies of the champions, the short stories, the films, and now Arcane. Skins that reference the story (whether canonical or alternate universe) acknowledge this dispersal. A skin such as Young Ryze (a very rare skin) literally depicts a champion at an earlier stage of life than the base game shows; it tacitly acknowledges that the full story of Ryze's youth is not told in the game — it lies in a missing narrative, which the skin only hints at. The young Ryze is thus presented as a trace of an unwritten prequel, a "missing origin" made semi-present. The "origin story" of *LoL* champions is eternally deferred; skins and external media combine to create an origin that is never fully present in any text, but emerges in the mediums between them.

By framing skins as traces, we see them as carriers of memory and continuity in the transmedial storyworld of *LoL*. They ensure that when the player switches from watching Arcane to playing *LoL*, they have a tangible reminder of the series in the game (and conversely, seeing the skin can prompt the player to watch the series). Skins function as "mnemonic devices" — visual reminders of the stories — and as bridges that carry emotions and knowledge across media boundaries. This capability greatly increases player engagement, as fans feel a sense of context and personal investment in the world across formats. A fan who loves the portrayal of a character in Arcane can continue this emotional engagement by using the Arcane skin in-game to weave their own interactive stories that reflect the narrative of the series. From a transmedia perspective, this is a powerful form of user-driven storytelling: the player uses the pieces (skins) provided by Riot to create experiential connections between the narratives of the game and those of other media.

However, it is important to note that the clues are inherently fragmentary and open to interpretation. Not all players will read skin's narrative clues in the same way. Some will ignore the story and just enjoy the aesthetics; others will fabricate their own fan theories that connect the skin to the story. This openness is intentional — Riot gives just enough information (say, a short skin description or a login screen comic) to ground the idea but lets the community fill in the blanks. This participation gap can be creative (facilitating fan fictions, discussions), but it

can also be perceived as the company exploiting the creativity of fans to enrich the IP at low cost.

5.1.3. Champion Skins as Pharmakon

At this point, I return to Stiegler, with whom I can examine technological artefacts in terms of their impact on human attention, memory and desire. He adapts the notion of pharmakon, which means both medicine and poison, to describe technologies (the technologies and the practices that surround them) as essentially pharmacological *(Philosophy Now* 140). That is, any technological system can be therapeutic in some ways and toxic in others. This framework is very appropriate to League of Legends skins, which can be seen as a kind of psychotechnology, or in Han Byeong-Chul's term, a tool of psychopolitics, designed to capture the attention of players and influence their behavior. Skins are both a "cure" — enhancing enjoyment, personal expression and connection to the world of play — and a potential poison — fueling compulsive spending, endless circles of engagement and commodification of play. In this section, I explore this duality, focusing on how looks externalize memory and modulate time and attention in the context of *LoL*'s seasonal content cycle.

Building on the work of Husserl and Derrida, Stiegler spoke of tertiary memory or tertiary preservation: externalized carriers of memory that store experiences outside the mind (from writing and art to digital media) (De Preester 105). According to him, video games and their content, for example, can be seen as tertiary memory devices — they store narrative and the players' stories. The skin is, in Stiegler's terms, the externalization of memory and identity. When a player acquires a skin that commemorates a special event, such as the Victorious skin, a player can receive for achieving a gold level in a season. This digital object becomes a memento of your achievement — a wearable trophy that reminds the player — and signals to others — the time and effort they put into the game that season. It effectively turns a piece of your autobiographical memory (your victory in Season X) into a lasting external relic in the game. It can be empowering and satisfying: Stiegler would say it's a cure for memory, as it enhances your ability to remember (the look will remind you of that victory long after you may have forgotten the details). It also adds layers of the player's personal narrative to the game world. Many players attach sentimental value to certain looks because of when or how they were acquired — looks become a repository of emotions and memories. In this way, skins function in what Stiegler calls mnemonic accessories that support human memory and experience outside the individual's mind (De Preester 105). They contribute to what he calls a technical milieu — an environment of accumulated tertiary recollection that shapes our

consciousness before we are even aware of it (De Preester 105). A new player who enters *LoL* today encounters a game that is already filled with these traces (years of skins, each with lore and player stories); his attention and imagination are shaped as much by this archive of skins as by the official narrative.

But pharmakon is a double-edged sword. The same skins that externalize memory can also erode certain abilities or lead to addiction. Stiegler has often warned of how contemporary digital products short-circuit attention and can create addictive feedback loops — a poison that "produces stupidity" if not translated into collective intelligence (De Preester 108). The *LoL* skin system is deeply intertwined with the monetization and engagement of gaming, and thus with what has been called the attention economy of digital media. Skins are a primary source of revenue (as they are cosmetic, they do not affect gameplay, which is consistent with fair competition, but it does mean that success depends on convincing players that they want these purely aesthetic objects).

Riot employs a number of tactics to keep players hooked and craving skins: timed events, rotating skin sales, loot boxes (Hextech crafting with random skin shards), prestige skins that require grinding or event pass purchases, etc. These design choices modify players' time — encouraging daily check-ins, longer playthroughs or participation in event windows — and focus their attention on in-game goals that are ultimately related to earning skins. For example, during RiotX's Arcane event, missions were placed so that playing the game unlocks Arcane skins on a staggered schedule. This meant that players had to log in each week to obtain a new free skin, effectively aligning player behavior with the release schedule of episodes in the series and marketing pressure. Here we clearly see what Stiegler might call the industrialization of attention: the game's technical system is fine-tuned to capture the attention and focus of players and direct it in ways that serve not only narrative enjoyment but also corporate goals (e.g. increasing the number of concurrent players during an episode, cross-promoting the show, and enticing some players to spend money on other skins while they are logged in).

The concept of pharmacology invites us to ask the question: do skins heal the relationship between players and the game world, or do they exploit it? The answer is inevitably both. On the one hand, many players express that skins increase their attachment to their favorite champions (you might like Ahri more after getting a beautiful skin to suit your tastes) and thus increase their overall satisfaction and willingness to continue playing — a positive, pleasure-giving effect. Skins can also stimulate creativity and social interaction: players can

have fun matching themed skins in a match, or talking in anticipation about upcoming skin releases, or creating fan art and cosplay based on skins. In Stiegler's terms, these are noetic outcomes — they contribute to "collective individuation," where community is formed around shared tertiary retentions (the skins and the aesthetic experiences they provide). The communal aspect can also be understood as an antidote to the alienation often experienced in online environments: it gives fans something positive to rally around and enrich their symbolic environment.

On the other hand, the forced cycle of acquiring skins can be problematic. Riot's strategy often takes advantage of FOMO by releasing exclusive skins or making skins available only through significant play (e.g. the Prestige system, where you either have to grind a lot or pay for extra loot tokens to unlock prestige skins for a limited time). Artificial scarcity and time pressure can create anxiety or unhealthy gaming habits — players feel they have to play for X hours or spend money, or they lose the chance. This can be seen as toxic engagement and fits what De Preester would call "dark patterns" of free-to-play design. Stiegler would probably classify them under the toxic aspect, where game technology "undermines the human psyche" by clinging to reward pathways and exploiting desire (De Preester 106). Indeed, gamers who generate revenue by buying skins or generate ad impressions through prolonged play can be seen as a form of "unpaid digital labor." From this political economy perspective, players are happy to work (play) for hours on end and actually generate value (through data, community vitality, direct purchases) for Riot. The carrot is often the desired look; the stick is the attention system designed to withhold or reward in carefully measured ways. The philosopher's question then becomes: does this pharmacological setup empower gamers or turn them into "comfortable digital workers" under the illusion of entertainment while monetizing them?

Stiegler's thoughts on the war for attention are prominent here. In a saturated media space, games compete for user attention. *LoL*'s frequent skin releases (usually a new batch every two weeks or so) and seasonal events (Lunar New Year, Halloween, Summer, etc., each with its own exclusive skin and mission) create a constant temporal rhythm that structures players' attention throughout the year. The game rarely lets the psyche rest; as soon as one event ends, the next one begins, each with its own cosmetic rewards and associated lore trinkets. This can be seen as *LoL* providing a beneficial service — there's always something new to enjoy, a living game world that keeps giving you content (the "antidote" to boredom). However, it can also lead to burnout or a sense of obligation when the game starts to feel like a second job to

keep up with the rewards (the poison of hyper-attention). Stiegler would probably suggest that care (cura) is needed in our engagement with play — that without careful consideration, these mechanisms can undermine our attention and turn play into a kind of automated consumption. Perhaps he would argue for what he calls noopolitics — a politics of attention that seeks to transform these psychotechnologies into something that actually benefits collective individuation rather than short-circuiting it (De Preester, 105).

Interestingly, Riot sometimes explicitly acknowledges the need to balance engagement and burnout when designing looks. For example, after community outcry over event fatigue, they have sometimes lowered mission requirements or given more free rewards to ensure goodwill. This suggests the idea that pharmakon is manageable: the dose makes the poison, as Paracelsus said. If skins and events are tied too aggressively to monetization, players will rebel (which hurts the longevity of the game). If done in moderation and with the players' enjoyment in mind, skins can really enrich the game. In line with Stiegler's thought, in a way the challenge is to turn toxic aspects into medicine for Riot to design in a way that cultivates long-term care and attention for the game as a cultural object, rather than extracting value. There have been instances that could be described as therapeutic uses of skins: for example, Riot sometimes gifts skins to players to apologize for downtime or problems, using cosmetics to alleviate community frustration. Skins also often come with charitable campaigns (e.g. skins where a portion of the proceeds go to fundraising organizations such as Extra Life or social impact funds), flipping the script to present the purchase as a contribution to a greater good rather than mere consumption. These illustrate the flexibility of pharmakon — the same mechanism (selling looks) can be framed as exploitative or altruistic depending on context and intent.

Finally, Stiegler's focus on memory and time allows us to see how looks modify players' experience of time in the game. All skins, especially the thematic skins tied to events, anchor the game in time — in the chronos of the franchise. Years later, when you look at a skin in your collection, you can be transported back to the time of its release (your personal chronology as a player). Skins actually help you build your personal gaming history. This can increase the importance of engagement (the game was not just an endless now; there are seasons and eras you lived through). But, conversely, there is a danger that if you are constantly focused on the next skin or the next season, you live in a state of constant partial attention, never fully enjoying the present before preparing for future content — a problem Stiegler suggests may be related to our culture losing its epoch (pause for reflection).

In summary, through Stiegler's pharmacological lens, *LoL* views appear as a double-edged technique. As medicine, they deepen the game's ability to delight, memorialize, and engage players in a rich world — externalizing memories and identities in reinforcing ways, and sustaining a long-term community through shared symbols. And as a poison, it can push players into a treadmill of attention and consumerism, where the intrinsic pleasure of play can be overshadowed by external rewards and status symbols. Recognizing both sides is essential for a balanced understanding. In the next chapter, I will explore how Riot's treatment of the cure in the service of world-building — using skins as a means of aesthetic and brand-building continuity for continuity across media — can be understood as an antidote to transmedia fragmentation.

5.1.4 Aesthetic and Brand Consistency

One notable achievement of League of Legends is the expanding franchise coherence. Despite the franchise moving into different mediums — a realistic art style animated TV series, K-pop music videos, comic books, and more — it has maintained a strong brand identity and aesthetic continuity that fans instantly recognize. Champion sketches play a strategic role in this consistency. They act as aesthetic stabilizers that align the look and feel of characters across different platforms and ensure that the world to be explored remains consistent, even when stories are told in different formats. In this section, I will attempt to explore how skins contribute to branding and narrative cohesion, focusing on the case of the Arcane and K/DA musical universes, and how they maintain artistic consistency. Following that, I will also discuss how this serves Riot's transmedia strategy — blurring the line between gaming and other forms of entertainment to enhance immersion and player loyalty.

When the animated series Arcane launches in 2021, the League of Legends will showcase its champions through a whole new medium and style. The series' art direction is sumptuous and painterly, more detailed than the cartoon-like graphics of the video game. Characters such as Vi, Jinx, Caitlyn and Jayce were reimagined with nuanced costuming and facial design to fit the television narrative. To bring this aesthetic back into the game, Riot introduced Arcane skins (for Jayce, Vi, Jinx, Caitlyn) that faithfully reproduce the characters' Netflix appearances in the game. This served several purposes. For one, it reassured fans that the Arcane versions of the characters are "canon" enough to be included as skins in the game's roster — a nod to the fact that the series' story is now part of the franchise's official lore (in fact, after the success of the series, Riot updated the lore of several champions in the game to align with Arcane). Second, it preserved visual continuity: the player can seamlessly transition from watching Arcane to playing the game as Arcane Vi, without any distracting differences in design. Coherence reinforces identification; the Vi you bonded with on Netflix is recognizably the same Vi you now control in Summoner's Rift, just rendered with the game engine. Thus, skins buffer the confusion that occurs in transmedia franchises when characters look or feel too different in different mediums. The result is a more immersive transmedia experience where the boundaries between watching and playing seem thinner. Riot has explicitly conceptualized Arcane as permeating all of their games and experiences (the RiotX Arcane event was literally

designed to “infuse the magic of Arcane into our games” and skins were the primary method for the actual *LoL*.

From a branding point of view, these cross-media skins impose a unified image. *LoL*’s brand is rooted in its champions — their personalities and iconic looks. When these champions enter a new medium, Riot must carefully manage their portrayal to stay true to the brand. Skins allow Riot to bring the look back into the brand’s core arsenal. For example, the Arcane Jinx skin ensured that Jinx’s neon-punk look from the series (with his rose-tinted glasses, for example) was part of Jinx’s official aesthetic roster, rather than a one-off design that only lives in the series. Likewise, the K/DA skins made what could have been a wild departure — *LoL* champions as pop stars in an alternate universe — an integral part of the *LoL* brand offering. By producing high-quality K/DA skins and performing the songs at *LoL* esports events, Riot signaled that K/DA was as much “League of Legends” as the story of the base game. This synergy benefits the strength of the brand: it means that all roads lead back to *LoL*. Whether it’s the K/DA music video on YouTube or Arcane on Netflix, the fan’s entry point is the K/DA music video, fans will find their way to the game and see that content reflected there, reinforcing the sense of a unified, cohesive franchise.

Skins also help to maintain artistic consistency. Riot’s art teams ensure that even skins inspired by external media follow the game’s art guidelines for model proportions, color vibrancy, etc. For example, the Arcane skins had to translate a more realistic style into the slightly exaggerated proportions of the *LoL* game models. The end result still has a “*LoL* feel”, which is key to not alienating players with aesthetic expectations. In addition, skins allow iterative improvements to character design in a way that can retroactively influence the brand. One notable example is Caitlyn: Arcane introduced Caitlyn with a slightly updated costume and rifle design, and around the same time Riot gave Caitlyn an Art & Sustainability Update (ASU) in-game with a new default costume (closer to Arcane’s style) and reworked all of her skins for consistency. Here a media feedback loop occurred: Arcane’s sophisticated aesthetic informed the game’s default aesthetic through an update, while Caitlyn’s Arcane skin provided one-on-one accuracy for the show. In effect, the skins can act as a testbed or a transitional tool for updating the overall art direction of the game. Because they are optional and numerous, introducing a new art concept through a skin is less risky than completely redesigning the default character. If it is popular (as Arcane was), this concept can be incorporated into the main look.

In terms of narrative coherence, the skins serve to maintain a thematic continuity between stories. Riot has created several alternate universes (AUs) for skins — e.g. Star Guardian (magical girl AU), PROJECT (cyberpunk AU), Spirit Blossom (mythic anime-inspired AU), etc. Each is essentially a pocket narrative or aesthetic theme that spans many champions. In a sense, this is transmedia within the game itself, but it goes beyond that (some AUs have been accompanied by comics and anime, such as an anime short for Star Guardians). By making these AUs internally consistent (each has its own logo, design language, sometimes its own music/themes), Riot treats them almost as sub-brands of *LoL*. Players can become fans not only of the champions, but also of the thematic look lines. This strategy means that even if the content of the game is wildly varied (high fantasy knights one moment, mecha robots or pop singers the next), it does not split the brand because they are all well labelled and coherently developed. It's like Marvel having different comic book series — diverse content under one umbrella. The looks serve as concrete products that anchor these underworlds. A player can identify *LoL* with bright, adventurous fantasy (if they like the Battle Academia and Star Guardian skins, which have a very anime-adventure vibe) or edgy sci-fi (if they like the PROJECT and Odyssey skins). But as they are all clearly part of the League skin universe, together they enrich the brand rather than dilute it. Riot reinforces this by cross-promoting the skins in other mediums: for example, merchandise such as figurines often depict the champions in popular skins, rather than just their basic look, further reinforcing these skins as part of the official iconography.

The esports dimension also benefits from an aesthetic continuity driven by skins. In *LoL*, each annual World Cup has a unique thematic identity (dragons in 2017, cyberpunk in 2018, etc.) and Riot often releases a championship skin that commemorates it with a unifying visual motif (blue and gold color scheme, etc.). Over time, the Championship skins form a series of skins that fans associate with the annual rhythm of the tournament. In addition, each year the winning team will work with Riot to design World Championship skins for the champions of their team members' choice. These skins will include the team logo and often the personal style cues of the professional player, yet they should look like *LoL* and appeal to a wider player base. It's a delicate balancing act: skins commemorate an esports moment (which can be seen as the "real world narrative" of the tournament) and include the team's branding, while also being a timeless part of the *LoL* skin catalogue. Riot has generally succeeded in doing this — the fan community looks forward to the winning skins each year and sees them as an honorable continuation of the game's story. By involving the players in the design and

giving them a share of the proceeds (the League of Legends World Champion gets his own skin as a reward, Riot strengthens the link between professional players and the game's content, effectively merging the competitive narrative with the in-game narrative. From a branding perspective, this means that the esports story is literally embedded in the product. When someone uses an SKT T1 skin on Zac (when SK Telecom T1 won the World Championship), they carry the flag of that championship with them to every match, subtly propagating Riot's esports brand through everyday gameplay.

All these examples point to one of the main functions of skins in transmedia: they weave a coherent fabric from different strands of content. They ensure that whether a story comes from a Netflix show, a music album or a championship stage, it can be aesthetically and narratively reinserted into the central universe. This gives players the feeling that everything matters — no part of the franchise is left disconnected. It also serves corporate interests by keeping engagement within the Riot ecosystem. When a fan falls in love with Jinx through Arcane, the Arcane Jinx skin draws them into the game. If a long-time player becomes interested in K-pop through K/DA, that interest will continue to be tied to *LoL*, rather than drifting towards, say, other K-pop groups not related to the game.

From a critical perspective, one could say that this is a form of brand control that ensures that Riot's IP remains self-referential and closed. By providing official transmedia skins, Riot is perhaps preventing fan-created content from becoming the dominant way of introducing external influences into the game. This is a way of saying "if Jinx becomes a pop star, it will be through our sanctioned K/DA skin and our music, not solely through fan art". In fairness, Riot does embrace fan creativity (for example, they often spotlight fan-made skin cosplays), but always in the shadow of the official skin lines. This tight control is typical of the media conglomerates that manage the franchises — similar to the way Disney ensures that Star Wars looks like Star Wars in the films, series and merchandise. The difference in games lies in the interactive sense of ownership: by buying or acquiring a skin, the player has a personal stake in the aesthetic of the franchise. In effect, they become ambassadors for the brand every time they use the skin publicly in a match.

On a positive note, this continuity allows players to connect to the story universe on their own terms. For example, a player can choose to dress up as a version of Vi Arcane and replay the series' fights, or they can start a full K/DA lookalike team and imagine a crossover scenario. The consistency of the design makes these imaginative exercises even more

immersive. This is transmedia play rather than mere consumption: the game becomes a sandbox where different pieces of story from different media can come together. We saw this when players organized Arcane-themed tournaments where everyone used Arcane skins, effectively blending TV narrative and gameplay as a fan-driven event. Such exercises highlight how a well-managed aesthetic continuum can stimulate a participatory culture (the fun of mixing and matching story elements) while still being aligned with the official world.

In conclusion, *LoL*'s use of skins for aesthetic and branding continuity exemplifies an integrative and stabilizing transmedia strategy. Skins function as a common visual language that translates narratives across media. They help maintain Arcane's narrative coherence across platforms by anchoring designs in the game and ensuring that ventures such as K/DA enhance rather than fragment the *LoL* universe. This strategy has contributed to the success of *LoL* as a multimedia franchise — Arcane, for example, attracted many new viewers to the lore and as the game mirrored the show through skins and updates, the conversion of viewers into players was likely to increase. The champions remained recognizable and likeable regardless of the medium, and this was largely due to the careful care taken in their appearance through the looks. The skins are therefore not just a revenue generator, but a narrative glue and branding tool, demonstrating the sophistication of Riot's transmedia approach to storytelling.

5.1.5 Critical Considerations: Participation and Control

Having painted a largely positive picture of *LoL* views as enriching transmedia objects, it is essential to look critically at the wider implications of the phenomenon. The interplay of culture and commerce in digital games raises questions about commodification, digital labor and the nature of participatory culture under corporate governance. In this chapter, I address the potential critiques and complexities: are gamers really co-authors of the transmedia narrative, or are they skillfully managed as consumers and unpaid marketers? Do skins reinforce or exploit participatory culture? How is corporate control manifested in the stories that are or are not told through skins? I explore these questions while taking a balanced approach — recognizing the pleasure and creativity that looks can bring, but not excluding the power dynamics.

The narrative and the commodification of the game: in *LoL*, every skin is also a product on the market. This dual status — narrative artifact and commodity — can lead to tensions. On the one hand, players value skins as added content; on the other hand, the aggressive promotion of skins and the sheer quantity (currently over a thousand) reflects a profound commodification of the game aesthetic. Some critics argue that modern games are less and less about providing the full experience upfront and more about creating the possibility for ongoing microtransactions. In the case of *LoL*, lore and storytelling elements through skins are often locked behind paywalls (exceptions such as free Arcane skins or event rewards). For example, the story of Star Guardian — which includes multiple lookouts and accompanying short stories — requires the purchase of champion skins to fully enjoy the thematic resonance of the game. While the stories can be read for free on a website, the story implementation (for example, featuring the Star Guardian Ahri) costs money. This raises the concept of "meta-paywalling" narrative immersion. The narrative is told not merely through the cut-scenes that everyone sees, but embedded in the objects of purchase. Riot has thus commodified the transmedia narrative to a certain extent: it monetizes the desire for expanded narrative content by making skins the delivery mechanism.

From a political economy perspective, this is an example of the convergence of the attention economy and the commodity economy in games. Criticizing views celebrating participatory culture, scholar Christian Fuchs notes that while corporations own and control platforms, user participation (even if enjoyable) is deeply intertwined with exploitation — users

produce value, In *LoL*, players "work" by constantly engaging with the game, hyping the skins, creating content (streams, fan art) around them, which increases the franchise's visibility and profits. It could be argued that the overflowing transmedia experience (fans enjoying the skins, cosplaying, etc.) masks a working relationship where players are effectively unpaid promoters. The notion of digital labor applies here too: players invest time (often huge amounts of money) in unlocking or presenting skins, which in turn promotes the game and encourages others to play or spend.

This does not mean that players get nothing — they get fun and social capital — but the exchange is unequal in terms of power and the distribution of profits. Ultimately, the economic decisions are made by Riot: pricing, availability, design decisions are all about revenue. A critical reader might ask, for example, whether certain popular skin themes have been expanded not out of narrative necessity but because they sell well. (If Star Guardian skins are hugely popular, Riot will continue to produce them even after the "story" may have narratively run its course; the story can always be extended to include more because the demand is there.) In this sense, narrative is subordinate to commodification. The notion of pharmakon mentioned earlier can be recalled here: the "medicine" of narrative depth is packaged as the "poison" of consumption — to get the full experience, one must keep buying.

Henry Jenkins praises fan participation - the idea that players and fans actively contribute to content (through fan art, fiction, mods, feedback) is a feature of modern media (3). The *LoL* community certainly exemplifies a participatory culture: fan-made skin art is widespread, theorizing about future skins or alternate universes is common on forums, and cosplayers bring skins to life at conventions. However, Jenkins also acknowledges (and Fuchs in a 2014 article specifically point out) that this participation takes place within channels that are often curated, or at least profited from, by the company (3). Riot is known for engaging with the community, soliciting skin ideas, and sometimes even incorporating memes or fan concepts into skins (for example, the fan concept of "Pizza Delivery Sivir" eventually became an official skin due to popular demand). This suggests a collaborative relationship where fans have a voice. This can be empowering: players feel heard when a silly meme skin becomes a reality, or when their cries for an Arcane skin are heard. Such feedback loops foster a sense of co-creation.

Yet we must ask: who sets the conditions for this participation? Riot ultimately filters out what is feasible or appropriate for the brand. It is unlikely that they will accept a fan idea that does not match their plans or that could damage the brand image. The participatory culture

around *LoL*, while vibrant, largely operates in a way that complements Riot's marketing. Fans who cosplay or create fan art of skins are essentially grassroots advertisers (albeit out of sincere affection). Riot encourages this with contests and spots, but it's a managed participation. It's far from a free-to-play game where fans can, say, directly contribute their own skins to the game (as opposed to a game like Dota 2, where user-created skins can be added via Steam Workshop — a more direct participatory content creation, though ultimately still curated by Valve). Riot maintains strict control: only their employees can design and publish skins, and players can only get access to them through official channels. This ensures revenue streams and brand consistency, but it also signals the limits of fan management. Fans can suggest and celebrate, but cannot self-publish within the game ecosystem.

Fuchs devotes a whole chapter in his volume, *Social Media: A Critical Introduction,* that Jenkins' idea of a participatory culture glosses over issues of hierarchy and ownership (76). League of Legends does show that fans are mobilized and integrated into the growth of the franchise, but on Riot's terms. Riot's skill is that it makes it feel fun and collaborative rather than exploitative. Many players probably do not feel "exploited" when they purchase a skin. They are happy with their purchase and their enthusiasm. This is in line with modern theories of power in media — control need not be experienced as oppression; it often works by aligning the interests of the company with the pleasure of the users. Despite this, some critical voices in the community question the pricing (e.g. the increase in $100 prestige skin packs at events was controversial) and the prioritization of cosmetics over improving game balance.

Another critical issue is representation and narrative control. Riot decides which stories to foreground and which to avoid through looks. For example, some champions rarely get story-rich skins, perhaps because they are not as popular (they have less profit interest), which means their story remains undeveloped. This may bias the narrative universe towards certain characters or themes at the expense of others. Furthermore, although *LoL* skin narratives explore a number of fantasy/sci-fi tropes, they tend to steer clear of overt political commentary or real-world social issues, so the transmedia story remains 'safe" and globally marketable. This is understandable for a game company, but it's a form of editorial control that can be disappointing for those who hoped games could deal more with serious issues. Even Arcane, while mature in tone, fits squarely into a dystopian steampunk fiction that doesn't directly challenge the status quo of our world (though it does have implicit messages about class differences).

Finally, consider the agency of players within the transmedia narrative. Looks allow players to express themselves (the choice of look is a form of self-representation). However, the possibilities of expression are predetermined by Riot. It is a postmodern form of collage of expression: players can combine elements (skins, emotions, champion choices) to create meaning, but cannot create entirely new elements. This contrasts with truly user-generated content platforms or moddable games, where players can introduce new designs. *LoL*'s closed system means that agency is about selection and performance, not about creating from scratch. It begs the question: does it matter? Many would argue that curating from a rich selection of content is sufficient for meaningful participation. Others might lament the lack of more democratized content creation (for example, why shouldn't players be able to design custom skins for personal use? The answer is probably that it undermines the business model and quality control).

As far as exploitation is concerned, it is worth noting that the players' enjoyment is real and often worth the money. However, exploitation can also be subtle: for example, the existence of gambling-like mechanics (random loot boxes for skin shards) can be seen as an exploitation of psychological weaknesses for profit. Riot has somewhat reduced its reliance on pure RNG loot compared to a few years ago, partly due to community resistance and regulation. This indicates that the voice of the community can to some extent control corporate practices. The cancellation of particularly unpopular monetization plans shows an overall play of power — players collectively have some power if they organize their dissent (we can recall EA's infamous Star Wars Battlefront II loot box fiasco as an industry parallel that led to change due to fan dissent). Similarly, the *LoL* community has long since stopped plans to sell power directly, consolidating an approach that only provides cosmetic revenue. The criticism now is not of pay-to-win (*LoL* avoids this) but of the creeping sense of pay-to-enjoy-cool-stuff.

In summary, the critical view highlights the balance between player agency and corporate governance in the transmedia ecosystem of *LoL*. The positive reading is that skins allow fans to actively participate in the franchise narrative, personalize their experiences and be part of the global creative community that celebrates the game. The critical reading sees a carefully orchestrated system where each expression of fan enthusiasm is fed into a monetized cycle, and where the company retains a decisive role in which narratives are amplified. Both readings are valid, and indeed they coexist. *League of Legends* has achieved a kind of symbiosis with its player base: a participatory culture that is genuine, yet structured by corporate

strategies. As media scholars, we can appreciate the art and communal joy of *LoL*'s transmedia storytelling, while acknowledging that *LoL* operates within a capitalist framework that exploits culture for profit (as do all the big media franchises, from Marvel to Harry Potter).

5.1.6 Conclusion

By deepening the academic analysis of *LoL* skins, I wanted to show that even highly commercialized digital artefacts can be read through the lens of media philosophy and game studies to gain insights into our contemporary media environment. One insight is the erosion of clear boundaries between media: a skin is both a piece of game content and a narrative puzzle, which can include a Netflix episode and a music video. This reflects the larger trend of a convergence culture, where the media industry seeks to maximize synergies and where audiences are fluidly following stories across platforms (Jenkins 95-96). *LoL* has pioneered this gaming space and its model is increasingly being followed by other live-action games that aspire to transmedia franchises.

Another important lesson is how player engagement is nurtured. *LoL* shows that players are invested not only in the success of the gameplay, but also in the cultural capital: they own a piece of the story, express their identity through skins, and interactively experience the stories they love. This suggests that game companies that invest in rich world-building and narrative integration (rather than treating skins as a separate cosmetic sale) are likely to achieve stronger and more lasting engagement. The success of Arcane (which won awards and attracted new fans) has fed back into engagement with the game and justified a strategy of building narrative depth rather than shallow attachments. For scholars, this underlines the importance of analyzing games not just as economic products, but as ludic-narrative texts that carry meanings and shape the subjectivity of players.

I also highlighted some critical challenges. The commodification of game content raises concerns about accessibility and equity — narrative enjoyment should not be limited to those who can pay, but in practice full participation in the transmedia narrative of *LoL* often is. Critiques of digital work and participatory culture remind us that fan enthusiasm can be put at the service of corporate goals. There is an implicit contract between Riot and its community: Riot provides the playground and cool games (like skins and crossovers), and the community provides the time, attention and money. As long as both feel they are getting value, the relationship is symbiotic. But it can become problematic if, for example, monetization strategies are overdone and fans feel exploited, or if creative directions conflict with fan expectations. Ongoing negotiation and feedback are therefore key; interestingly, Riot actively engages with player feedback (via social media, PBE test servers, etc.), which shows that it recognizes that in a live service, community goodwill is as important a currency as dollars.

Ultimately, League of Legends and its epic cinemas show that 21st century transmedia world-building is not just about transmitting a story, but about creating an ecosystem where consumers become participants, where story is integrated with product, and where the boundaries between play, viewing and creation are blurred. This offers exciting opportunities for engagement and creativity, but also warns caution about the conditions of engagement. As games continue to evolve as narrative platforms, the lessons of *LoL* — both its success in creating a rich, enjoyable transmedia universe and its critique of monetization and control — will be instructive for designers, scholars and players alike. The humble "skin," once a mere texture change, has now become a pillar on which much of the new media paradigm balances — indeed a digital object that deserves serious attention.

5.2 Semantic Databases and Cultural Agency: The Hungarian Database of English Studies

With the advent of data in the humanities — the transformation of rich cultural texts and research findings into digital data — contemporary philology is facing a radical transformation. Scholars warn that this process, driven by large-scale digitization and computer analysis, could reduce subtle nuances of interpretation to mere statistics. For example, Alberto Romele, already discussed in more detail in previous chapters, diagnoses a digital habitus characterized by the "algorithmicizing" and "big datafication" of cultural life. At the same time, digital philology projects aim to make scientific use of computer methods. These efforts must find ways to reconcile the often positivist logic of digital tools with the interpretative ethos of traditional philology. In this chapter, I present a case study of a new digital humanities project — the Hungarian Database of English Studies (HADES) — which aims to address this challenge. Drawing on the debates around data hermeneutics and digital positivism, I present the plans for HADES as a semantic bibliographic database for Hungarian English Studies, potentially also using artificial intelligence. The aim of my chapter is to explore the possibilities for integrating recent digital infrastructures and long traditions of interpretation into a coherent framework, and what this might mean for knowledge production and dissemination in the humanities.

The field of digital humanities (DH) is a fundamental example of this interdisciplinary challenge. Broadly speaking, DH can be seen as the intersection of traditional humanities (e.g. literature, linguistics and cultural history) and computational methods and technologies. Digital tools such as text mining, network analysis or spatial analysis allow humanities scholars to collect, archive and analyze cultural data on a previously unimaginable scale. This new field offers new perspectives: large-scale text analysis can reveal historical patterns of language use, and digital archives can democratize access to rare manuscripts. However, scholars of critical data studies, as demonstrated earlier in my dissertation, argue that the process of transforming culture into data itself raises a number of meaty questions. If research in the humanities becomes reducible to numbers and algorithms, how can the interpretative richness of texts be preserved? Conversely, if philologists reject digital methods as "positivist," they risk ignoring valuable evidence. Romele stresses that the challenge is not only technical but also ontological: the digital age has shaped a new worldview that is reshaping meaning-making itself (p. 2022). To navigate this change, we need a "data hermeneutics" — a methodological approach that applies

hermeneutic inquiry to digital data, rather than unquestioningly accepting computational results.

5.2.1 Digital Datafication and Data Hermeneutics

At this point, I would like to return to the analysis of Romele's digital hermeneutics, which highlights two fundamental problems. According to Romele, the collection, analysis and trading of user and consumer data has led to a "general algorithmization" and "big data-ization" of social reality (148). More specifically, it is about automated processes that track and encode our behavior, creating a data superstructure that frames culture and knowledge. This phenomenon is already apparent in research: search engines, social media and digital libraries generate vast amounts of information (metadata, click-through data, text corpora) that can be statistically analyzed. For humanities scholars trained in close reading and contextualization, these types of quantified data matrices may seem alien. Romele warns that the resulting "digital habitus" shapes our perceptions and even our thought processes in almost imperceptible ways, often unconsciously (153).

Second, Romele, following in the tradition of Ricoeur, insists that hermeneutics should evolve accordingly. Traditional hermeneutics treats texts (especially sacred or classical works) as linguistic objects to be interpreted by readers. In the digital age, however, our objects are often not just printed pages, but data sets or algorithmic processes. A purely quantitative approach — a kind of "digital positivism" — would analyze word frequencies, network graphs or data visualizations without paying special attention to the meaning of the data. To counter this, critical theorists propose a data hermeneutic that combines traditional and digital methodologies. The term "data hermeneutics", coined by Gerbaudo in digital media studies, refers to a method for interpreting the "deep structures of meaning" inherent in digital corpora (97). Gerbaudo argues that the advent of Big Data has brought a quantitative bias to the social sciences: models and metrics have proliferated, but issues of context, nuance and human intent are often ignored (98-99). He argues that to fill this gap, methods of interpretation need to be adapted to digital formats: for example, selecting subsets of social media posts for close reading, or treating individual posts as data points and part of a larger discourse (105). The data hermeneutic approach thus calls for a resistance against the positivist ideology of "dataism" and for the restoration of qualitative interpretations in digital research, inherently close to the humanities.

In philology, this debate has a deep resonance. Classical philological work (textual editing, philological commentary, literary history) is fundamentally and inescapably interpretative. It observes and evaluates authorial intent, historical context, stylistic nuances.

In contrast, some digital projects ask, in effect: how many times is word X used in corpus Y, or what is the network of references in these texts? Such indicators may indeed reveal new patterns, but they do not answer why a literary motif is repeated or what its cultural significance is. As Gerbaudo, for example, notes, an over-emphasis on "data analysis" can marginalize meaning (96). At the same time, it is also an important consideration that the discipline of philology could stagnate if it rejects the possibilities of digital analysis, research such as the discovery of forgotten texts through OCR or the visual mapping of networks of influence spanning several centuries. The challenge for HADES is therefore symbolic: to design a digital infrastructure that can capture rich metadata and network connections while allowing traditional humanities disciplines to interpret these connections.

5.2.2 The HADES Project: Digital Tools, Traditional Philology

The aim of the Hungarian Database of English Studies (HADES) is to collect and organize the scientific results of Hungarian English Studies on an interconnected digital platform. By using semantic technologies, HADES aims to achieve the transformation of the various bibliographic records into machine-readable knowledge graphs. This includes the linking of authors and institutions, publications and keywords, possibly full texts and translations. In this way, HADES is building a specific, interconnected open data ecosystem for English Studies in Hungary.

The reasons for creating the HADES are interdisciplinary and pragmatic. On the one hand, the field of Hungarian English Studies is fragmented between universities, journals and archives. Traditional bibliographies, in print or solid form, are quickly becoming outdated and difficult to search. For example, it may be difficult for a specialist to find all references to "Byron" in Hungarian academic literature before 2000, or to identify who has worked on lexical semantics in Hungarian English Studies. In contrast, a semantics platform can support complex queries (e.g. 'show all articles on Byron by ELTE authors") and visualize collaborative networks. Knowledge graphs — interconnected structures of concepts and metadata — are designed to support discovery, retrieval and navigation in scientific domains (Haslhofer et al. 2). By representing entities (people, texts, ideas) and the relationships between them, knowledge graphs transform archives into navigable "global knowledge networks" (Haslhofer et al. 2). HADES aims to apply this model so that, for example, an AI-driven search can easily link an author's name to his or her publications, the topics he or she has written about, and the references to his or her works. In this way, HADES aims to be both 'social value" (making science publicly accessible) and future-proof, "AI-enabled" (structured enough to be suitable for machine reasoning).

Technically, we plan to base HADES on the semantic web standards. Its metadata model will most likely use the Resource Description Framework (RDF) standard to encode information about individual bibliographic items as triples (subject-item-subject). Ontologies or vocabularies from the Linked Data community (e.g. Dublin Core for titles and authors, BIBO for bibliographic records, FOAF for persons and SKOS for subject headings) will provide a common schema. For example, an article can be encoded with RDF properties that link its title to its publisher, its author to a URI that identifies the scholar, and subject terms to standardized

concepts. Such structuring means that the data can be queried using SPARQL and that external datasets (e.g. ORCID or Crossref) can be integrated using common identifiers. In the case of HADES, this results in a digital infrastructure where bibliographic knowledge is represented explicitly. This will not only enable sophisticated search and analysis by scholars, but also allow downstream AI applications to make use of the data — for example, in the future AI agents could automatically generate summaries of Hungarian Byron studies by reviewing the graph.

Equally important is the interdisciplinary cooperation that underpins the HADES. Building a semantic-based portal requires the collaboration of programmers, librarians and philologists. Philologists provide the domain knowledge: they define which entities are important (authors, works, genres, translations), what relationships exist (mentor-mentee relationships, reference networks, thematic groupings), and how to make names explicit (e.g. Librarians and information scientists bring standards of metadata and data curation, ensuring consistency in recording publication dates, journal titles and authoritative sources. Information scientists will implement the back-end system: creating the triple repository, providing scalable performance for and designing the user interface. Such teamwork must cross institutional boundaries. Research in the field of English Studies is ongoing in several universities, such as ELTE, PPCU, DE and HUN-REN research institutes. Each institution may have its own repository or bibliography. The success of the HADES depends heavily on convincing these stakeholders to make data available.

If we look only at the results of contemporary English studies, permission may be needed from a department to share publication lists and republish data. Likewise, departmental norms differ: a history department may prefer to evaluate annotated catalogues, while an English literature department may focus on individual literary studies. As opposed to MTMT, where data structures are based on affiliation, person and maybe larger categories of disciplines, HADES would extensively focus on the boundaries of English Studies in Hungary. These cultural differences may also present methodological challenges. For example, to align metadata, there needs to be agreement on vocabulary: should an entry be labelled “Victorian literature” or “19th century British novel”? Overcoming such obstacles in the project itself requires a hermeneutic approach — a recognition that the curation of infrastructure is an interpretive activity.

Finally, the HADES must reconcile qualitative reporting with quantitative effectiveness. As Gerbaudo noted, digital research runs the risk of valuing the easily quantifiable over the

meaningful. To avoid this, the design of HADES counts on an interpretive layer: each bibliographic record has space for texts or notes, and controlled subject terms are taken from established classification systems in history, literature and linguistics. These features ensure that researchers or academics using the database can find not only title lists but also contextual information about why a certain work is significant. In summary, HADES seeks to follow the insights of digital hermeneutics: it uses quantitative methods without abandoning its humanistic commitment to context and meaning.

5.2.3 Case Study: Byron's Reception in Hungary

In the most optimistic scenario, the Hungarian Database of English Studies will eventually cover all areas of English Studies, but we need to start somewhere, at a tangible scale. A particularly striking example for the creation of a database is the reception of Byron in Hungary. Lord Byron's poetry has had a complex afterlife in Hungarian literature: it has been revered as a symbol of romantic freedom, translated by poets of several generations, and invoked in nationalist and even rebellious discourses. Until now, however, the bibliographical record of its reception has been scattered or incomplete. In 2024, Miklós Péti, associate professor at the Institute of English Studies at Károli Gáspár University, discovered a manuscript long thought to be lost: the complete Hungarian translation of Byron's *Childe Harold's Pilgrimage*, produced in 1904 by the young Dezső Kosztolányi. Kosztolányi (1885-1936), although one of the most important figures in 20th century Hungarian literature, was a writer and, alongside his work as a cologne poet, one of the best-known translators. Yet, until this discovery, only fragments of his Byron translation were known. Péti identified the text in archival documents and cross-referenced it with the entry list for the Kisfaludy Society translation competition. The find is significant for the history of Hungarian literature because, it is "the only complete translation of this work [Childe Harold] by a major 20th century Hungarian poet." (Péti 345). In other words, Kosztolányi's translation of Byron is not only interesting, but also a unique cultural monument that establishes a link between English Romanticism and Hungarian modernism.

In the context of the HADES, this finding also illustrates the promise and necessity of semantic bibliographies. Once Kosztolányi's translation is in the database, it can be linked to several entities: the original Byron poem, the Kisfaludy Society competition, Kosztolányi's authorship, and 1904. Such links were virtually invisible in traditional catalogues — which is why the discovery of this translation is so significant. Importantly, HADES may also be able to record metadata about the discovery in the form of an event/action (e.g. Péti's identification in 2024 and the consequent publication in the journal *Vigília*). Thus, a researcher can search in HADES for the phrase "Byron's works translated by Hungarians" and immediately get Kosztolányi's entry alongside other entries (published or to be published). Furthermore, the public interface of HADES can highlight this as news: a blog or timeline feature can announce "Rediscovered: Kosztolányi's translation of Byron (1904)", making the scholarly results available to a wider audience. This illustrates, among other things, HADES's public service

goal: a footnote to a philological study can become a useful resource for teachers, students, and even AI-based cultural applications.

The Byron example also clearly points to an interpretive caution: cataloguing such a translation requires carefully prepared decisions. Do we simply put it in the category of "Byron reception" or do we link it to themes such as romantic nationalism? How do we note its literary significance? HADES designers need to incorporate curatorial, and possibly editorial, decisions into the data model. These decisions will determine how future readers (whether human or machine) will interpret the entry. HADES can thus be not only a passive repository, but an active player in structuring the knowledge about Byron in Hungary. This reminds us that digital bibliography is itself a kind of literary criticism, which requires a hermeneutic knowledge of its own codes.

5.2.4 Technical and Methodological Infrastructure

The HADES database uses modern semantic technologies to achieve its goals. It is built on the paradigm of linked data, with RDF (Resource Description Framework) as its backbone. In practice, each bibliographic item is represented as a set of RDF triples (e.g. "Kosztolányi-authorOf→Childe_Harold_Translation_1904"). RDF allows linking data from different categories: authors and institutions, texts and translations, etc. Ontologies and schemas (e.g. Dublin Core for bibliographic attributes, FOAF for personal data, BIBO for reference information) provide a common vocabulary. By conforming to these standards, the HADES database will be interoperable with other semantic resources — for example, an author's URI can be linked to his ORCID or even to his records in the OSSK.

This model is supported by a triple storage database and a SPARQL endpoint. Queries can be made through the graph to extract information. For example, a SPARQL query can be used to retrieve all the works written by Hungarian scholars about Byron, or English-Hungarian translations written between 1900 and 1950. The use of OWL (Web Ontology Language) provides the possibility to define richer class hierarchies (e.g. "Poet → Hungarian poet → Translation") and inference rules. These semantic features are essential for creating machine-readable meaning. In the case of HADES, this means that bibliographic data can be used for visualization or analysis purposes in AI tools, or can be used by other digital humanities projects around the world.

For users, HADES will probably offer search and browsing functions enhanced with semantic filters (faceted search by author, subject, date, etc.). It can also display visualizations of the knowledge graph, such as co-author networks or publication timelines. Importantly, HADES supports both peer-reviewed and public use. Experts (philologists, librarians) can download RDF data or use the SPARQL interface for more complex research. The general public and academics can use a web site or portal with a user-friendly interface: keyword search, interactive browsing and plain-language entity descriptions. This dual focus also solves the problem of data hermeneutics in practice: professionals can do a "close reading" of the datasets, while non-professionals get an accessible, narrative-based scientific presentation.

Behind the system is a workflow integrating library information science and digital infrastructure. The bibliographic data can be drawn from existing sources: library catalogues, publisher metadata, national databases and scientific biographies. HADES needs to collect and

normalize this data, a process similar to ETL (extract-transform-load) in data engineering. Ideally, we should be able to perform automatic data queries from library database APIs; human editors then check for accuracy (e.g. For content tagging, we select or adapt literary ontologies (e.g. a controlled list of genres, literary trends or critical themes). Each step requires interpretation: deciding whether a given book is "romantic poetry" or just "19th century literature" will affect how users find it.

However, there are also institutional challenges. Hungarian universities have different practices regarding open data. Some institutions may restrict data sharing or not have the possibility to digitize old documents. A convincing demonstration of the scientific and cultural value of HADES requires proof of its social utility. The Byron discovery itself can serve this purpose: it shows that HADES records important scientific developments as they occur. In addition, by aligning with international trends, such as Linked Open Data for cultural heritage, HADES coordinators can ensure support from funding organizations investing in digital sciences.

Finally, there are methodological questions about combining approaches. HADES is not simply a library of books and articles, but also an analytical platform. For example, if HADES integrates full-text data (e.g. from digital editions), it can facilitate text mining. However, this raises the risk warned by Romele and Gerbaudo: quantitative content analysis without meaning. The HADES documentation should therefore provide explanations for interpreting the data. For example, if the HADES provides the number of keywords related to Byron in Hungarian journals over time, the description should point out that the numbers alone do not give a picture of the critical context (e.g. did interest increase or decrease? Which aspects of Byron were discussed?). Thus, HADES can take a kind of data hermeneutic stance: it provides powerful new tools but also encourages users to read the data with a critical eye.

5.2.5 The Implications of Digital Hermeneutics for the History of the Field

The creation of the semantic infrastructure of HADES actively transforms the contours of Hungarian English Studies, and indeed of philological knowledge. By deciding which entities and relations to include in the data model, the project puts a theory into practice. For example, the foregrounding of certain subject terms or the linking of certain national literatures will determine the image of a given field. These design decisions are acts of interpretation. A digital ontology can embed a hierarchy that separates linguistic studies from literary studies, or it can merge them into broader cultural studies. Either way, researchers browsing the HADES will perceive the topology of the discipline in the way the designers intended. In other words, the architecture of HADES both reflects and constructs scientific categories. This is a major responsibility.

Moreover, the shift to semantic data also brings certain epistemological issues to the fore. Traditionally, philologists have archived knowledge in prose, in textual form: a bibliographical entry may contain notes on the reliability of the source or on the narrative context. In HADES, knowledge is divided into discrete, formal units. This has the advantage of accurate retrieval ("find all of Lord Byron's translations by year"), but risks losing narrative nuance. As a result, HADES users have to deal with a new kind of interpretation: meta-interpretation of data. They will ask why a work is categorized under certain themes or why certain relationships exist. This reflects Foucault's view of archaeology: the organizing principles of the archive determine what counts as historical truth. The semantic layer of HADES is a modern archive of knowledge; it invites critical reflection on how data structures shape understanding.

In an adherence hermeneutic sense, HADES is both a hermeneutic object and a tool. The project has to confront "dataisms," i.e. the belief that the available data exhaust the meaning. If, for example, the HADES search engine ranks results by number of citations, it may favor frequently cited authors and marginalize new voices. Designers should therefore build in interpretive safeguards: allow users to filter by qualitative criteria (e.g. peer-review status, subject area, language of publication) and encourage skepticism about numerical indicators. The HADES may also include explanatory notes or digital scientific publications alongside the entries to link the data to textual interpretation on an ongoing basis.

More broadly, HADES is an example of how technology can mediate knowledge. The semantic network of Hungarian English Studies is itself a kind of hermeneutic framework: scholars learn about their field through the categories of HADES. According to Romele, the task is to create the conditions for "distancing" from the digital habitus, i.e. users need to be aware of how digital constructs guide perception (158). The ultimate value of HADES will be judged not only by its completeness, but by whether it enables users to process data critically. If it succeeds, it will strike a balance: leveraging AI-enabled semantics to expand access and analytical possibilities while preserving the interpretive core of philology.

5.2.6. Conclusion

The Hungarian Database of English Studies illustrates the critical development of the humanities. From a technical point of view, it shows how semantic web technologies (RDF, SPARQL, ontologies) can be applied to build a discipline-specific knowledge graph. Conceptually, it responds to the epistemic tension between quantitative digital methods and qualitative literary interpretation. By focusing on "data hermeneutics" — the idea that digital infrastructures themselves require interpretation — HADES joins scholars such as Gerbaudo and Romele in calling for a hermeneutic turn in the era of big data. Furthermore, the interpretive implications of HADES remind us that digital transformation is never neutral. Every schema and data point in the platform is full of choices about which knowledge we consider important. Precisely by making these choices explicit and retrievable, a semantic infrastructure like HADES encourages users to reflect on them. It makes visible the "deep structures of meaning" that underpin our science, allowing us to extend hermeneutics to the digital world. In this sense, HADES points to a new paradigm: one in which knowledge in the humanities is not only produced by human scientists, but also shaped by the architecture of information. Such projects do not diminish the importance of interpretation, but translate it into a data and networking language, prompting experts to develop a "digital philology" with a critical perspective for the 21st century.

5.3 Cultural Techniques of Print and Digitality: What Shakespeare's First Folio And ChatGPT (Not) Have In Common[1]

The debut of OpenAI's ChatGPT made the autumn of 2022 a pivotal moment when AI text generation became a mainstream experience. The novelty comes not from the existence of such AI-based tools, as simpler chatbots have existed for decades, but the fervor and fascination they sparked. ChatGPT and similar models are capable of producing fluent, contextually mostly appropriate language on almost any topic, highlighting a potential that has captured the imaginations of millions of users and prompted tech giants to race these tools into their popular platforms. In the process, these AI systems have become deeply woven into the fabric of the "platform society," wherein our communication and knowledge practices are channeled through large corporate platform infrastructures as theorized by van Dijck.

The public reception of generative AI has been quite heavily polarized. Opinions on the two sides oscillate between utopian hopes and apocalyptic fears. The initial hype surrounding ChatGPT made civil discussions almost impossible, with some declaring the greatest event of our lifetime and others visioning the end of the world as we know it. One side hails AI-based tools as a revolutionary piece of technology that changes everything. They urged immediate deployment to all walks of life: using AI-based technology as productivity booster, creative collaborator, as friend, coach and even mental health advisor. The other side quickly dismissed the new tools due to unreliable outputs, ethical concerns, and the upcoming replacement of human workers. For every successful use-case, everyday experience showed a simple failure: an elegant paragraph of analysis may turn out to be factually baseless, a harmless query might provoke biased or absurd answers. The strong reaction of the general public reflects genuine uncertainties: although it is not yet clear whether this technology will prove a transformative infrastructure or a passing fad, it certainly appears to be venturing into new territories of human experience.

From an academic perspective, these debates signal that something fundamental is at stake. The advent of large language models once again invites us to examine how meaning and

[1] Originally published in Patkós, Gábor. 2024. "Cultural Techniques of Print and the Digital: What Shakespeare's First Folio and ChatGPT (Not) Have in Common." In *Artificial Intelligence, Digital Literacy, Digital Pedagogy*, 2-9.

knowledge production is under reconfiguration in a data-driven culture. The very fact that a machine can appear to produce coherent, context-sensitive language without any human understanding has profound consequences on the ontology of meaning. ChatGPT serves as a profound example, illustrating the concept of datafied ontology I explored earlier in my dissertation: it is trained on billions of words – essentially human knowledge captured as data – and engages with users through statistical pattern-matching rather than through any human-like understanding. In an editorial regarding the launch of GPT3, Floridi somewhat skeptically observes that we have effectively "decoupled the ability to act successfully from the need to be intelligent, understand, reflect, or grasp anything", paving the way instead to "agency without intelligence" that characterizes these models ("AI as Agency", 14-15). This development leads to serious epistemic question: how should we understand the status of texts generated by an algorithm that does not create knowledge through understanding or intuition, but synthesis and prediction?

Shifting epistemics leads further uncomfortable questions about the nature of truth. If the majority of modern thought organized knowledge around reasoned discourse, the present condition of digitality organizes it around information streams. Referring again to earlier chapters, in 'infocracy', the information regime of 21st century, according to Han truth becomes structurally precarious: time is fragmented, discourse is displaced by the circulation of data, and generative AI accelerates a shift from shared facts to personalized informational outputs. In this view, the very facts, as Arendt puts it, which "assert themselves by being stubborn, and their fragility is oddly combined with great resilience" (259) is under transformation by the dominance of information. As digitality makes everything reproducible and quantifiable, Han comes to the conclusion, that "the digital is directly opposed to factuality" (50). This shift marks a clear fracture from the tradition of modernity, the primacy of facts and rational discourse.

Both the First Folio and contemporary AI platforms mark moments when new media systems altered how language circulates and how culture defines authoritative knowledge. In both cases, technological innovation (the printing press in one, machine learning in the other) interacted with institutional power (the early modern publishing industry and patronage networks, versus today's tech corporations and data economies) to produce new possibilities for discourse. By placing these cases side by side, I aim to explore how cultural techniques of inscription, interpretation, and transmission change through medial shifts. The chapter will thus move from analyzing the digital ontology and hermeneutics of AI back to the print-historical moment of the First Folio, using each as a mirror for the other. Through this juxtaposition, I

would like to provide insight into the long arc of mediated knowledge: how the information machines of different eras (a printed folio in the seventeenth century, a generative language model in the twenty-first) manage to reconfigure the relationship between human beings and the language that binds their culture.

5.3.1 Discourses of Print and AI Platforms

At this point I need to circle back to Kittler's concept of discourse networks, which he defined as historically specific configurations of media techniques, institutions, and practices that organize the production and circulation of knowledge. It is arguable whether today's AI platforms can be seen as a new discourse network. The idea seems straightforward: a configuration in which neural network algorithms, cloud infrastructures, training datasets, content moderators, and user prompt practices together form a system for producing and distributing discourse. On the other hand, it seems to be giving too much credit to AI platforms by attributing the power of developing new discourses to them. I rather find it to be an outcome of the discourse network of digitality already raging.

Looking at large language models through this lens, we see them as the latest stage in the evolution of the cultural techniques of digital data operations. The printing press invented the technique of mechanically reproducing texts, broadcast media centralized the dissemination of audiovisual information, the AI platform brings a technique of prediction: producing novel-seeming content from statistical digestion of prior human texts. The materiality at play here is not paper or electromagnetic waves, but code, servers, and data. Yet the effect, much as Kittler would predict, is a reorganization of how knowledge is stored, communicated, and valued. The datafied ontology of our time means that who we are, what we know, and how we decide are increasingly mediated by algorithmic interpretations of data.

These questions highlight the cultural and epistemological disruption provoked by AI platforms, however, are just illustrations of a larger epistemic shift. They operate in what Nick Couldry calls an emerging "anti-hermeneutic" paradigm, which I briefly addressed in the opening section. According to Couldry, in a world of algorithmic power, data-centric systems bypass traditional interpretation; they pattern-match rather than read, optimize engagement rather than communicate meaning, which is primarily a tendency of Big Data and AI. Basically, a mechanism to short-circuit meaning, this mode of operation must be critically addressed by renewed hermeneutics attuned to our algorithmic age. The task then once again leans into debates on digital hermeneutics, which seeks to adapt interpretative methods to the anti-hermeneutic paradigms of machine-generated content and algorithmic filtering and understand how meaning is being processed by computational systems, and what that means for human knowledge.

Therefore, it is important to consider large language models not just as isolated algorithms but as socio-technical systems, or more accurately as platformized information machines. Their creation and operation depend on platform infrastructures: colossal datasets scraped from online content, cloud computing power concentrated in big tech companies, and deployment via interfaces that millions use. The cultural techniques at play here include datafication, algorithmic training and tuning, and the conversational interface that structures how users query and receive answers. I argue that these digital techniques can be considered as extensions of age-old techniques of information management: classification, retrieval, synthesis, and generation of text.

At first glance, Shakespeare's First Folio, a collection of Renaissance plays in print may appear unrelated to 21st century AI-based technology. But the First Folio was itself a pivotal media artifact that reorganized cultural knowledge through new technological and institutional practices. Published seven years after Shakespeare's death, the Folio was the first comprehensive assemblage of his plays, preserving eighteen dramas that had never been printed before and arranging them in a systematic format (it even introduced the division of plays into comedies, histories, and tragedies). In effect, it transformed ephemeral stage performances into enduring literature, helping to canonize Shakespeare's works as central to English heritage. The Folio's compilation and publication were the product of a cultural technique of early modern print: collecting manuscripts, editing texts, and reproducing them via the printing press to reach a wider public. This process solidified certain texts as authoritative and worth preserving, thus shaping the trajectory of literary culture and scholarly interpretation for centuries to come. By comparing the media functions of the First Folio with those of AI text platforms, we can illuminate what is unique and what is perennial in the way media technologies mediate knowledge.

5.3.2 Cultural Techniques of Print

It would be tempting to simply try to place digital culture somewhere on the axis of oral culture and literacy. The recognition of the similarities between post-typographic culture and primary oral culture is reflected in Walter J. Ong's concept of 'secondary orality", which he defines as "the individualized introversion of the age of writing, print, and rationalism which intervened between it and primary orality and which remains as part of us" (*Rhetoric, Romance, and Technology* 285). The recognition itself is very important, and I will return to it, but terminology that extrapolates one aspect of digital technologies can easily lead to oversimplification. Ong's other work, *Orality and Literacy* deals with the psychodynamics of verbosity and the relationship between subject, spoken and written language is a remarkable scholarly achievement and a type of investigation that seems promising for evaluating digital culture, rather than narrowly focusing on the notion of secondary verbosity. Therefore, in order to understand how First Folio and large language models such as ChatGPT continue to shape our relationship to texts and culture, we need to analyze the characteristics of text production technologies, independently of the content they produce. As N. Katherine Hayles puts it, "the different technologies of text production suggest different models of meaning; [...] they create new kinds of text worlds" ("Virtual Bodies" 69).

McLuhan is one of the first and most prominent thinkers to develop media- or technology-oriented research. Medium Theory is a research methodology or medium-specific critique that focuses on "the specific characteristics of particular media or types of media" (Meyrowitz 50). In contrast to media studies that focus specifically on the content of different types of media, media theorists are interested in characteristics that are distinct from content. However, an even more important and perhaps more difficult question to answer is where we start to consider something as a medium.

The technological invention of the 15th century, printing with movable type, fundamentally changed European thinking and brought the triumph of literacy over oral tradition, as highlighted in the work of *Orality and Literacy* by Ong. However, the emergence of new technologies never in itself represents a major change. We can only speak of a new paradigm when a piece of technology becomes widespread and starts to develop its own apparatuses or networks of discourse. Macho, in a seminal work, introduces the concept of cultural technologies that can be used to illustrate this distinction:

> Cultural techniques — such as writing, reading, painting, counting, making music — are always older than the concepts they are derived from. People were writing long before they had any idea of writing or the alphabet; it took thousands of years for pictures and sculptures to create the concept of the image; and people still sing and play music today without knowing anything about sound or musical notation. Counting is also older than the concept of numbers. Of course, most cultures counted or performed some mathematical operations, but this did not necessarily give rise to the concept of number (179, qtd. in Winthrop-Young 8)

Even after the invention of printing, oral and written culture largely co-existed, but it was only when printing itself became a medium that it began to have a real impact on the production of knowledge and the structure of culture. A more familiar example of this phenomenon is the invention of the smartphone. The first iPhone was largely seen as a mobile phone whose features were seen by many as an unnecessary luxury. People mostly used it as a traditional mobile phone with extra features, but as the features and capabilities of each model expanded, the "extras" became more commonplace. Eventually the smartphone became a device, a medium in its own right, that fundamentally changed the way we interact with our environment. Likewise, the invention of printing led to the spread of literacy, the diversity of ideas, the faster spread of information and the emergence of copyright, as explored for example in Febvre and Martin.

The philological instability of Shakespeare's texts is a fundamental knowledge of Shakespeare scholarship, and often the subject of exhaustive research. In fact, relatively little is known about Shakespeare's life or even his manuscript, especially compared to the impact of his works (see e.g. Greenblatt 2004, 2010). Many of Shakespeare's plays are common compositions, with Gary Taylor suggesting that "less than two-thirds of Shakespeare's plays are Shakespeare plays" (2). Only about half of the plays were published in print during the author's lifetime, mostly in quarto format, according to Dawson and Kennedy-Skipton for example. Some plays were altered, or even exist in separate versions. The best known example is *King Lear*, which was published in quarto (1608) and folio (1623), while editors created their own texts based on one or the other version. In fact, what we know about Renaissance theatre in general is that it operated in a rather unstable context.

Analyzing the First Folio as a key example of the cultural technique of print, it can be argued that one of the key drivers of Shakespeare's lasting impact and the image of artistic

genius today was the success with which Shakespeare's colleagues used book printing. Of course, these claims are not exclusive to the First Folio. They can also be said of books in general, but in this article I examine the issue through the example of one of the world's most famous books, the First Folio. Without going into the extremely complex but fascinating history of Shakespeare's reception, I will justify my claim as follows:

Firstly, printing allowed Shakespeare's plays to be effectively distributed to a wider, more heterogeneous audience, dispersed in space and time. The novelty of writing, and printing more broadly, is that it exists in material form, unlike the spoken word. As Ong perceptively notes, it is impossible to stop and hear sound at the same time. Once the sound stops, we hear only silence. In other words, "sound exists only when it ceases" (*Orality and Literacy* 31). Writing, on the other hand, is "forever tyrannically locked into visual space" (*Orality and Literacy* 11). Unlike watching and listening to actors perform on stage, readers can immerse themselves in the written text of plays on demand. Rather than having to recite the text of the play, the printed text allows them to make the play their own — or at least the written version of it.

Secondly, it led to the development of a cult of the artist. While theatrical productions are by their nature collaborative and necessarily require the work of a company, written text can easily be assumed to be the work of a single author. Of course, given the laborious process of printing, this is never true of the book itself, but the idea fits well with the authorship of the text. Disregarding the source of the text in question (collaboration? translation? adaptation?) or the material problems of the text, the possession of an authentic volume provides the authority needed to establish the text's ownership.

Thirdly, the First Folio functioned as an effective archive. It provided the subtitles and stability needed to create a reference collection of texts. As Ong argues "A text cannot be directly refuted. Even after a complete and devastating refutation, it says exactly the same thing as before" (*Orality and Literacy* 78). As we saw earlier, one of the essential features of Shakespearean philology is that manuscripts are lost. Performances also mostly fall outside the category of archives, as there are no records of many aspects of the staging of the plays. The first folio solved this problem by becoming the first authoritative source from which to 'study" or "look up" (both words derived from literary culture) Shakespeare's plays.

At this point, it seems appropriate to conclude that the first folio, as an early example of culture entering the Gutenberg galaxy, strongly influenced the development of the Shakespeare canon and fundamentally shapes our general view of literary texts. However, in the post-print, information network world, our relationship to texts, authorship and originality seems to be changing again. With the continuous development of information technologies and the availability of computing capacity to store more and more knowledge production and cultural expressions in the digital space, the 2020s increasingly look like the decade of the emergence of AI-based technologies.

5.3.3 Knowledge production and AI-based Platforms

First of all, it is important to clarify that the term AI, especially in the context of large language models, is not without its own problems. Since there are serious and in-depth debates about the interpretative or cognitive processes involved in the operation of large language models, the use of intelligence in the human sense is problematic. Instead, the way these models work can be more accurately described by the term prediction, based on the analysis of statistical relationships. Manovich in his work on the aesthetics of artificial intelligence, stresses that "if we want to better understand the difference between visual media synthesis methods of artificial intelligence and other representational methods developed throughout human history, then the use of the concept of prediction, and thus the description of these artificial intelligence systems as predictive media, is a good description of this difference" ("Seven arguments" 12). Therefore, the task of human understanding and knowledge production falls entirely on the reader or user of the system.

Interestingly, the way that generative AI technology (such as LLMs) works is similar to the world of pre-print oral culture, which can be compared to a certain extent to Renaissance theatre. First, the absence of original works. For the outputs of the model, there is neither a clear, verifiable original work nor a manuscript. Instead, we have so-called "black boxes", or machine-learning models, to which we are referred, which suggests that it is difficult to explain how the model arrived at its conclusion. In most cases we know what goes in and what comes out. Even if we know exactly the principles of how the algorithms work in a given model, and can thus explain the output, it is almost impossible to determine the precise processes behind the complex mechanisms that drive LLMs (Savage no pag.). In short, we have to rely on assumptions, circumstantial evidence and second-hand information.

Secondly, the unreliability of the text. Another important similarity between the understanding of LLMs and the study of Renaissance theatre is that in both cases we have to take into account an inherent philological instability. Responses to the same prompt, even using the same models, can vary significantly and are therefore ultimately definitive and to some extent unreproducible. In the case of a book, we are visually reading and trying to interpret the written text. However, when interacting with LLMs, we usually do this using natural language prompts, but as we have seen, these actions are mostly associated with a high degree of instability and randomness. In this sense, prompts in a situational context can be seen as

individual manifestations, similar to Bakhtin's trans linguistic theory of language (see Holquist on Bakhtin's Trans-lingustics). Not to mention that different versions of the same model may significantly alter the outcome of the same prompts, the internal logic of text generation to which we are accustomed, or we may have to consider a completely new data set on which the LLM was trained. As a result, we have to accept an innate fluidity that is very different from our educated minds.

Finally, we can conclude that generative AI as a cultural technique is very different from printed texts in terms of its archiving capabilities. Although by design LLMs store huge amounts of data continuously based on their training on unimaginably large datasets, they are clearly not built for information retrieval, like an encyclopedia or a database. Databases contain semantic information in the form of text, and we rely on text or text-based search to find information, whether it is a printed encyclopedia (index) or a digital database (search). Interestingly, LLMs are in some ways more like the structure of the human brain. Information exists in semantic categories, labels, relationships, networks and patterns, so in order to access or read it, you need to perform a very specific set of operations. However, there is simply no way to check that the output matches the data set used for training. So, if I ask ChatGPT to list all of Shakespeare's plays, I get an answer that may or may not actually be correct. I can only find that out if I look it up. However, I will not know, even after looking it up, what ChatGPT was taught, how it arrived at the number, how much my answer was influenced by my question, and how the answer will change the next time I ask it again. This is the true pharmacology (in Stiegler's sense) of generative AI technologies: the first foil crystallized, solidified and stabilized knowledge production and cultural expression, while generative AI returns it to its former unstable, contingent and uncontrollable state.

On the other hand, although the process of interacting with LLMs involves a high degree of randomness, and recreating the exact process that led to the production of the information is almost impossible, the output of the information itself is infinitely reproducible and infinitely extensible. We can command the model to expand the output by repeating the prompt as many times as we want, or the human agent can change the prompt and the text output. This paradoxical situation is explored by Hayles in her influential essay *Virtual Bodies and Flickering* Sign*ifiers*, in which he examines how the long-standing presence/absence dichotomy is being replaced by the pattern/absence dichotomy in the spread of information technologies in global digital culture societies (70-72). In the digital textual world, characterized by a degree

of reintroduced instability, or in Ong's term, secondary verbosity, the marker is not a unique marker but "a flexible chain of markers linked by arbitrary relationships determined by relevant codes" (Hayles 77). Digital cultural techniques evoke an earlier, less text-dominated period in which accessibility and availability were increasingly emphasized. As a consequence, we are entering a cultural condition where meaning is characterized by the instability of oral culture inscribed in the immaterial virtual body of the printed text.

2.3.4 Infocracy and the Crisis of Truth: The Role of AI-based Platforms

The transformation outlined at the previous section can be reframed by Han's account as a political shift from democracy to infocracy: power accrues to those who command the infrastructures of collection, storage, ranking, and delivery, ultimately the cultural techniques of digitality. In the condition of infocracy, social phenomena and public discourse is determined less by deliberation than by the logistics of information. Han identifies two main reasons that explain this development:

Firstly, the condition of digitality leads to temporal fragmentation. In his reading, information is momentary: it lacks extension, the duration needed for judgment and memory to congeal (*Infocracy* 21). In Han's terms, the constant flood of information flood erodes the temporal stability on which "knowledge, experience, knowledge, and insight" depend (*Infocracy* 21). Without the solid structures of facts, that stabilizes perception, an information-driven discourse becomes increasingly fractured and short-sighted until discourse itself falls apart. What cannot persist cannot be argued with; it is merely replaced by the next update.

Second, digitality leads to the disintegration of discourse. According to Han, discourse presupposes friction: the practice of rationality is a time intensive practice (*Infocracy* 21). Being able to contest premises, to test reasons, to revise claims requires longer term attention and thinking, frequently identified as a distinguished feature of print culture. Contrary to this, feed architecture minimizes friction. Algorithmic governance rewards and thrives on novelty and affective charge, not argumentative force. This is the essential operation of attention economy: when discourse yields to circulation, authority migrates from reasons to reach. The consequence of such public discourse is that it does not matter whether a claim is warranted but whether it is delivered and consumed.

We can clearly observe this in the practical operation of social media, for example: frictionless design choices are not merely cultural aesthetics; they are governance mechanisms. Recommendation systems, inherently part of any social media platform, exemplifies this observation while operating as a cultural technique: they materialize priorities (optimize for watch time), define publics (target segments), and distribute salience (ranked lists). The net effect is a political epistemology in which legitimacy is silently tethered to visibility metrics instead.

Arendt's distinction of two types of truths clarifies what Han claims to be lost. Arendt separates rational truth (a priori, demonstrative) from factual truth (contingent, worldly) (321). She further insists, that political freedom requires a background of facts that actors agree to leave intact (154-155). Which means opinions can vary only if accepted facts are based on common ground. In an infocracy, however, the background dissolves. As I outlined, personalized feeds disaggregate publics, and the logistics of attention normalize the co-presence of mutually canceling "facts." What appears as pluralism is often parcellation: each segment receives a different background, and public argument loses the shared referents that make disagreement meaningful.

Crucially, Arendt stresses the importance of politics respecting facts, the "stubborn things" that ground the world in common. The specific threat of digitality is not simply misinformation, as opinion or philosophical truth still accepts the common ground of factual truth. According to Han, it is the structural neutrality of the information regime toward truth as such (47). Discourse in the regime of infocracy simply "loses all reference to facts and factual truth" (47). Following the previous example of social media's governance algorithms, I would say, if systems are optimized for engagement, truth appears only as a side effect when it happens to align with engagement.

The effects of social media have been studied and discussed for a long time now. This is evident in Han's account of the infocracy. However, the addition of AI-based technologies to the operations of social media platforms, or generative AI platforms themselves intensify these dynamics along three axes relevant to truth:

Firstly, text and images can be produced faster than human vetting cycles can constrain them. The question is not whether a model "lies" but that it is non-responsive to truth conditions by design. It simply optimizes next-token likelihood under distributional constraints. The discourse emphasizes "hallucination," (utilizing the very human act of personification, interestingly) but that pathology is not a bug: it is what happens when statistical adequacy substitutes for evidential adequacy.

Second, retrieval-augmented generation and user-conditioned prompting tailor outputs to inferred preferences by design. The result is informational comfort: rather than confronting recalcitrant facts, users receive stylistically and perspectivally aligned answers. It is an important distinction then, that generative AI delivers information (in the form of digital data), not truth. It serves corporate interests by aligning format and framing with data-driven profiles,

not by submitting claims to adversarial testing. At best it is the burden of the user to check, reference and understand the information provided.

Third, even where models cite sources, the path from prompt to output is rarely reconstructable in reasons, only in weights, at best. I referred to this phenomenon earlier by the black box analogy. Explanations are themselves generated performances. This opacity compounds Arendt's concern. In simple terms, if we cannot trace the warranting chain that stabilizes facts, public acceptance becomes a matter of trust or acceptance of infrastructure, not evidence accessible to inspection or reason.

Based on Han's reading of Arendt, I would say that the crisis is not that people suddenly hate truth, nor that falsehoods newly dominate. Han emphasizes that the nature of crisis is that the background conditions that historically allowed claims to be adjudicated as true or false—temporal continuity, discursive arenas with shared premises, and institutions that bind reasons to consequences—are undermined by an order that treats all content as equivalent packets routed for attention (*Infocracy* 51). Truth becomes underdetermined by design.

Based on Han's analysis, I would propose three practical consequences of the widespread use of AI platforms, which are again, not novel consequences but highlight a new level of epistemic uncertainty:

First, because information is quantitative by definition, it is infinitely copiable, recombinable, and fabricable, agents rationally adopt suspicion as a default stance (*Inforcracy* 51). In such a world, assertions require extraordinary meta-work—provenance, verification, method disclosure—before they can even enter argument. This meta-work adds further unbearable burden on top of the burnout experienced due to the cognitive load on our psyche by the information regime.

Second, it leads to the crisis of narration (*Infocracy* 52). Lacking trusted common premises, communities stabilize meaning through narratives: coherent stories that provide closure amid informational noise. Narratives, however, can bypass the discipline of facts while preserving the feeling of explanation. This is the terrain on which "alternative facts", conspiracy theories and other forms of narration neutral to truth can thrive (*Infocracy* 52).

Third, when governance depends on dashboards rather than deliberation, interventions fix what is measured, instead of what matters. As a result we experience the datafied worldview. In a neo-postivist environment championed by "Big Tech" corporations, metrics colonize

judgment. Optimization replaces justification. I referred to this ideology as an anti-hermeneutic worldview in Couldry's words.

2.3.5 Hermeneutics for an anti-hermeneutic age

The question then, as I see it, loops back to a hermeneutic of the anti-hermeneutic age. Putting the determinist, inevitable argument aside, it is clear that digitality cannot be reversed. If it did, the effects would stay with us. The question then, is what makes a culture truth-capable under its constraints? To attempt to outline an answer to this question, I will list a few suggestions for the preconditions for the possibility of truth claims in a networked order.

First, the most pressing issue of temporality. Platforms and public institutions must create durational spaces in which claims persist long enough to be examined. What I mean here is the possibility to spend time with information. Examples would be archival affordances, that privilege updated records over ephemeral posts. Or a similar direction would be interface commitments to show the revision history of salient claims. To illustrate this let us compare the default interface of X or the traditional interfaces of blogs. Categories, layouts, hierarchies. Blogs are relics of the influence of print culture overflowing into digitality. X on the other hand an artifact of digitality. No hierarchy, no temporality, just endless information streams.

Second, machine outputs should carry portable, machine-readable context that names the training scope, any retrieval sources, the time and settings of generation, and subsequent human or automated interventions. This step would be able to effectively grasp and attest to the fact that data is never neutral, but in fact "cooked" (Gitelman 3). The aim is not just credibility, but contestability, so others could reconstruct enough of the pipeline to test and reproduce claims. When full disclosure risks privacy or IP, publish signed attestations and verifiable ranges rather than nothing. Of course, if the training of models were carried out in bad faith, it is not in the interest of corporations to be transparent about their models.

Third, truth needs institutions that bind words to the world, not just content policies. What it means is that public investment should support open, versioned datasets in high-stakes domains, such as elections, public health, climate, so claims can be checked against stable references. This in fact has been the case in print cultures for a long time, with careful book-keeping even in totalitarian systems. Independent adversarial review bodies must operate outside platforms and repeat producers of refutable harms should face meaningful sanctions.

Fourth, in civic contexts, personalization should be off by default or bounded by symmetry so that queries about public facts draw from the same evidentiary base across profiles. The most obvious areas would be health records, insurance policies and services for

basic needs as customary in most developed countries. So, for example a health insurer should not be able to offer a coverage advertising a lack of vaccination based on the political opinion of the potential client. When divergence is unavoidable, systems should disclose it and explain why it occurs. In that case personalization could in fact become an explicit, revocable choice rather than an invisible baseline.

Last, accuracy alone is too thin a standard for domains where truth matters. The most obvious example would be standard safety guidelines, such as instructions to design explosives or catering to ideologically motivated policies (for example: ChatGPT, why should we eat the rich?). AI platforms should be judged on their ability to deliver palpable reasons: traceable warrants, clear assumptions, and calibrated uncertainty rather than the performance of confidence or people-pleasing. Evaluation should reward answers that show how they know and what would change their conclusion.

To that end, I would assert that an AI-skeptical posture is not technophobic. It is a refusal to mistake information abundance for epistemic sufficiency. In Han's terms, the task is to re-introduce the stability of temporality into an order optimized for instantaneity and delivery. In Arendt's, it is to defend the pre-political fact without which politics degenerates into the administration of appetites. Nothing less is at stake than whether a society of networks can remain a society.

2.3.5 Conclusion

In this case study, I have examined two prominent examples of new cultural techniques that ultimately lead to paradigm shifts, both with significant implications for literary studies, texts and culture as a whole. Both the First Folio and the ChatGPT mark the transition from one era relying on one mode of notation to another, characterized by the changing material and media conditions of the time. The arbitrary and unstable nature of Shakespearean "texts" in Renaissance theatre in some ways resembles the basic modes of operation of LLMs. Nevertheless, print has left a lasting imprint on our psychodynamics, especially in terms of notions of authorship and originality, philological stability and a sense of archival preservation that are unlikely to change without ongoing cultural residues.

I also attempted to argue that under today's condition of digitality—what Byung-Chul Han calls infocracy—truth becomes structurally precarious: time is fragmented, discourse yields to the logistics of information, and rankings and reach displace reasons. Through Byung-Chul Han's reading of Hannah Arendt, I would also maintain that political freedom depends on a stubborn layer of factual truth we agree to leave intact. Yet, the information flood and personalized distribution dissolve this common world into a fragmented digital reality leading to a crisis of public discourse. For Han, information's momentariness and metricized visibility erode the durations judgment requires; for Arendt, the loss of shared facts hollows the forum where claims can be adjudicated. The crisis, then, is not mere misinformation but the systematic underdetermination of truth in systems optimized for circulation. To keep a truth-capable public thinkable, I attempted to propose temporal repair (durational spaces and visible revision histories), provenance by default, firebreaks against civic personalization, and institutions that bind words to worlds. As it would be impossible to restore a pre-digital past, the goal is to make common action possible again.

Given that today's society lives in a global, postmodern, digital culture—or, to put it more simply, a post-print culture — literary studies could benefit greatly from a multidisciplinary approach that includes material, technological and media realities in the study of literary texts in the broadest sense. We are, after all, all products of the cultural techniques we are used to in our own time. As András Kiséry wrote: "our theoretical choices often arise from a material or media situation from which we really cannot "reflect outwards", to theorize our way out, like a theoretical Munchhausen. And since we are already in it, let's do something

about it" (1083). In a cultural environment that can rightly be called an information flood, the humanities have the task of doing what they do best: practicing the solid and rich intellectual tradition of interpreting and understanding texts. Not because they have aesthetic or moral value, but because they are a fundamental expression of humanity.

6. Conclusion

In my doctoral dissertation I attempted to outline a comprehensive framework for understanding the profound changes engendered by the process of digitalization and datafication, arguing that these phenomena constitute the ruling episteme of the 21st century, which I called digitality. The core task I undertook is the construction and demonstration of a theory of digitality focused on interpretation, analyzing its underlying assumptions, mechanisms, and consequences for culture, knowledge production, and the human subject.

In the first chapter of my dissertation, I focused on establishing the conceptual framework of digitality. My central argument in this section was that the phenomenon of digitalization and datafication should be understood as a dynamic discourse network in the Kittlerian sense. I provided an overview of major theories of technology, culture and their intersection. I further argued that today's discourse network of digitality, unlike previous information systems, is fundamentally built upon the process of datafication. We can imagine data without digitality, but we cannot imagine the digital without the underlying binary data. The digital way of being, or the datafied ontology as I called it dictates that only that which can be transformed into discrete data and expressed in bits can truly exist, be stored, or be processed within this discourse network.

Crucially, throughout my analysis I relied on the critical humanities perspective that data is never merely "raw" or a "given," Instead, data is a cultural product, chosen, measured, and coded within a human framework. The work by Alberto Romele, focusing on digital hermeneutics, emphasizes that data are traces or records of reality, indexical representations that require active interpretation to transform into usable knowledge. This interpretation is necessary because data are semantically open, carrying no inherent truth value until contextualized.

I further posited that the operations of the data layer display striking similarities with perspectives of the philosophy of language. Following Austin and Derrida, I demonstrated that digital data is inherently performative. Furthermore, based on Hamacher's work I attempted to apply the concept of afformativity to describe data's unique structure of deferred potential and infinite iterability. Data promises meaning and utility, but this meaning is postponed until the moment of use, resisting complete consumption by any single application.

The second major pillar of my dissertation was the argument that the datafication of digital life should be approached not as a neutral technical output, but as a set of cultural techniques specific to the 21st century. The chapter titled Digital Data Operations as Cultural Techniques explored the fundamental operations, such as encoding, measuring, sorting, storing, and processing, that define the cultural order by operationalizing basic distinctions. Bernhard Siegert's work shows how cultural techniques discretize the world, turning continuous phenomena into countable, operable units, a logic that culminates in digital data. Sybille Krämer adds that digitality is rooted in the "emergence of operational symbolism," where abstraction allows symbols, like the alphabet or bits to be manipulated independently of immediate semantic meaning.

In the contemporary digital landscape, these techniques underpin a new type of epistemology. Knowledge is increasingly modular and recombinable, taking the form of lists and sets, such as databases rather than continuous narratives. This shift has tied into the debate around Big Data, epitomized by the idea that "correlation will replace causation" and theory will become obsolete. The critical conclusion that I proposed then, however, is that while massive data sets offer unprecedented scale, human insight remains essential to contextualize, verify, and explain patterns: theory is not abolished but complemented, reinforcing the indispensable role of the humanities.

Drawing on the philosophy of Bernard Stiegler, my intention was to integrate the technical operation of data with the co-evolution of the human subject. Stiegler argues that techné and episteme are inseparable and co-create human existence. In turn, I identified digital data as the latest form of tertiary preservation, memory stored externally, a process Stiegler calls epiphylogenesis. Data centers, algorithms, and digital archives embody a significant part of human memory, fundamentally shaping the consciousness and social reality into which individuals are born.

In my last chapter, the theoretical framework – encompassing digitality as an episteme, data as cultural technique, and the role of technology – is further illustrated by practical case studies that highlight the critical productivity of a humanities approach to digital culture. My focus was identifying the datafied ontology of three cultural phenomena inherently digital and cultural expressions that would not exist without digitality. Following the methodological considerations of case study research, I selected three specific examples that inherently demonstrate the cultural techniques of digitality at play:

1. League of Legends Champion Skins: These are analyzed as paradigmatic digital objects (Hui), defined by complex data, metadata, and relational networks. They function as Derridean traces, fragments of narrative and memory that connect the game to external media (like the Arcane series or K/DA music). Critically, the skins exemplify the pharmakon: they are powerful tools for narrative immersion and community building (medicine) but simultaneously anchor sophisticated monetization and attention-manipulation schemes (poison). They reveal the tensions of participatory culture operating under strong corporate control and commodification, where narrative is often subordinate to economic strategy.

2. The Hungarian Database of Englis: HADES serves as a case study for integrating digital infrastructure with philological tradition. HADES aims to transform disparate bibliographic records into a semantic knowledge graph using digital ontologies like RDF and SPARQL. This process requires making explicit design and classification choices, which are inherently interpretive acts that actively shape the topology and understanding of the discipline (the history of Hungarian English Studies). The project demonstrates the necessity of data hermeneutics in practice: the data structure itself becomes a reflexive object of interpretation, requiring scholars to critically reflect on how the archive's architecture constructs knowledge.

3. Shakespeare's First Folio and ChatGPT: By juxtaposing print and digital text production technologies, I my analysis I explore the long-arc of mediated knowledge. The First Folio, as a crucial example of print culture, functioned as a stabilizing cultural technique, establishing authorship, philological stability, and a reliable archive, fundamentally shaping literary culture. In contrast, AI-platforms such as ChatGPT, operating through statistical prediction rather than human understanding, reintroduce an inherent philological and epistemic instability. The fluid, unreliable, and unreproducible nature of LLM outputs shares striking similarities with the instability of pre-print oral culture and Renaissance theatre texts. I hope that the comparison underscores that digitality does not merely replace literacy but highlights the episteme of digitality characterized by the pattern/randomness dichotomy and an inherent fluidity that contrasts sharply with the archival stability of print.

Finally, I would like to reiterate the key points of my dissertation. The comprehensive examination of digitality culminates in an urgent call for the humanities to redefine their role in the data-saturated world. The field needs to change with the changing episteme around us. However, this does not mean that the humanities need to stop doing what it does best. In fact, the traditional tools of literary studies, philology, and critical interpretation are not rendered

obsolete; rather, they are the disciplines best suited to confront and interpret the cultural text of data itself.

The prevailing trend of neo-positivism and data-driven analysis threatens to marginalize interpretation, defining everything as an autonomous, "given" object. This development is clearly reflected in the ongoing crisis that we experience in our field. The humanities therefore must resist this by practicing the "solid and rich intellectual tradition of interpreting and understanding texts", but adopting an open approach to actual, present-day cultural phenomena. This requires developing methodological approaches like data hermeneutics, ensuring that quantitative scale is always complemented by qualitative context. The challenge is to integrate critical reflection, context, and interpretation with analysis, recognizing that the architecture of information systems profoundly shapes what we consider knowledge and truth. By theorizing digital data operations as a specific set of cultural techniques and engaging critically with the resulting datafied ontology, the humanities can move beyond a defeatist discourse and seize the opportunity for intellectual renewal, ensuring that the human element: values, meaning, and interpretation is not lost in the endless flow of bits and algorithms.

Works Cited

Albert, Reka, and Albert-Laszlo Barabasi. "Statistical Mechanics of Complex Networks." *Reviews of Modern Physics*, vol. 74, no. 1, Jan. 2002, pp. 47–97. *arXiv.org*, https://doi.org/10.1103/RevModPhys.74.47.

Alghail, A. a. A., and O. H. A. Mahfoodh. "Academic Reading Difficulties Encountered by International Graduate Students in a Malaysian University." *Issues in Educational Research*, vol. 26, no. 3, Aug. 2016, pp. 369–86.

Almási, Zsolt. "The Theatrical Fate of Shakespeare's Great Tragedies in Hungary : Metadata and Their Value." *Theatralia*, no. 1, 2024, pp. 145–66. *DOI.org (Crossref)*, https://doi.org/10.5817/TY2024-1-7.

Anderson, Chris. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*, 2008. Tags. *www.wired.com*, https://www.wired.com/2008/06/pb-theory/. Accessed 15. Oct. 2025.

Antonijević, Smiljana. *Amongst Digital Humanists*. Palgrave Macmillan US, 2015. *DOI.org (Crossref)*, https://doi.org/10.1057/9781137484185.

Anuar, Nadia, and Gurnam Sidhu. "Critical Reading Skills: A Survey of Postgraduate Students' Perspective of Critical Reading." *Pertanika Journal of Social Science and Humanities*, vol. 25, Jan. 2017, pp. 163–72.

Arendt, Hannah. *Between Past and Future*. With Jerome Kohn, Penguin Publishing Group, 2006.

Aristotle. *Aristotle: Politics*. Translated by H. Rackham, Later printing, Edition, Harvard University Press, 1932.

Austin, J. L. *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955*. 2. ed., [Repr.], Harvard Univ. Press, 2009.

Bachimont, Bruno. “Between Formats and Data: When Communication Becomes Recording.” *Towards a Philosophy of Digital Media*, edited by Alberto Romele and Enrico Terrone, Springer International Publishing, 2018, pp. 13–30. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-319-75759-9.

Balogh, Gergő. “Afformatív : A Performativitás Kritikája Werner Hamacher Munkásságában.” *Bia Hangja : Az Erőszak Irodalmi És Nyelvelméleti Reprezentációi*, 2021, pp. 79–105. *DOI.org (Crossref)*, https://doi.org/10.17048/Biahangja.2021.79.

Bjorn, Genevive, et al. “Reading Research for Writing: Co-Constructing Core Skills Using Primary Literature.” *Impacting Education: Journal on Transforming Professional Practice*, vol. 7, no. 1, 2022. *scite.ai*, https://doi.org/10.5195/ie.2022.237.

Bluemink, Matt. “Bernard Stiegler (1952-2020).” *Philosophy Now*, 2020, https://philosophynow.org/issues/140/Bernard_Stiegler_1952-2020. Accessed 15. Oct. 2025.

Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things out: Classification and Its Consequences*. 1. paperback ed., 8. print, MIT Press, 2008. Inside Technology.

boyd, danah, and Kate Crawford. “Six Provocations for Big Data.” *SSRN Electronic Journal*, 2011. *DOI.org (Crossref)*, https://doi.org/10.2139/ssrn.1926431.

Braidotti, Rosi. “A Theoretical Framework for the Critical Posthumanities.” *Theory, Culture & Society*, vol. 36, no. 6, Nov. 2019, pp. 31–61. *DOI.org (Crossref)*, https://doi.org/10.1177/0263276418771486.

Brentari, Carlo. “From the Hiatus Model to the Diffuse Discontinuities: A Turning Point in Human-Animal Studies.” *Biosemiotics*, vol. 11, no. 3, Dec. 2018, pp. 331–45. *Springer Link*, https://doi.org/10.1007/s12304-018-9329-8.

Casemajor, Nathalie. “Digital Materialisms: Frameworks for Digital Media Studies.” *Westminster Papers in Communication and Culture*, vol. 10, no. 1, no. 1, Sept. 2015.

www.westminsterpapers.org, https://doi.org/10.16997/wpcc.209.

Castells, Manuel. *The Rise of the Network Society*. 2. ed. with a new preface, [Reprint], Wiley-Blackwell, 2011. The Information Age / Manuel Castells 1.

Couldry, Nick. "Hermeneutics for an Anti-Hermeneutic Age: What the Legacy of Jesús Martín-Barbero Means Today." *Media, Culture & Society*, vol. 46, no. 3, Apr. 2024, pp. 659–67. *DOI.org (Crossref)*, https://doi.org/10.1177/01634437231217176.

Couldry, Nick, and Ulises A. Mejias. "Making Data Colonialism Liveable: How Might Data's Social Order Be Regulated?" *Internet Policy Review*, vol. 8, no. 2, June 2019, pp. 1–16. *DOI.org (Crossref)*, https://doi.org/10.14763/2019.2.1411.

Cumming, Alister, et al. "Students' Writing from Sources for Academic Purposes: A Synthesis of Recent Research." *Journal of English for Academic Purposes*, vol. 23, Sept. 2016, pp. 47–58. *ScienceDirect*, https://doi.org/10.1016/j.jeap.2016.06.002.

D'Alessandris, Francesca. "Alberto Romele, Digital Hermeneutics (London/New York: Routledge, 2019)." *Études Ricoeuriennes / Ricoeur Studies*, vol. 11, no. 1, July 2020, pp. 173–76. *DOI.org (Crossref)*, https://doi.org/10.5195/errs.2020.492.

Dawson, Gilles E., and Laetitia Kennedy-Skipton. *Elizabethan Handwriting, 1500-1650: A Manual*. Norton, 1966.

De Preester, Helena. "Life Is What You Fill Your Attention with – The War for Attention and the Role of Digital Technology in the Work of Bernard Stiegler." *Phenomenology & Mind*, no. 20, 2021, pp. 102–16. *DOI.org (Crossref)*, https://doi.org/10.17454/pam-2009.

Derrida, Jacques. *Monolingualism of the Other, or, The Prosthesis of Origin*. With Internet Archive, Stanford, Calif. : Stanford University Press, 1998. *Internet Archive*, http://archive.org/details/monolingualismof0000derr.

———. *Of Grammatology*. Corrected ed, Johns Hopkins University Press, 1998.

———. “Signature, Event, Context.” *Limited Inc*, translated by Samuel Weber, 10. paperback print, Northwestern Univ. Press, 2008, pp. 1–23.

———. “The Rhetoric of Drugs.” *Points...*, Stanford University Press, 2022, pp. 228–54. *www.degruyterbrill.com*, https://www.degruyterbrill.com/document/doi/10.1515/9781503622425-016/html.

Deutsche Forschungsgemeinschaft. *Digitaler Wandel in Den Wissenschaften*. Oct. 2020. *DOI.org (Datacite)*, https://doi.org/10.5281/ZENODO.4191345.

“DIGITAL PRAXIS IN THE HUMANITIES: A DH@YALE CONVERSATION SERIES.” *Whitney Humanities Center*, https://whc.yale.edu/program/digital-praxis-humanities-dhyale-conversation-series. Accessed 15. Oct. 2025.

Drucker, Johanna. “Humanities Approaches to Graphical Display.” *Digital Humanities Quarterly*, vol. 5, no. 1, 2011. *Digital Humanities Quarterly*, https://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html.

Febvre, Lucien, and Henri-Jean Martin. *The Coming of the Book: The Impact of Printing 1450-1800*. Verso, 1997.

Feenberg, Andrew. *Transforming Technology: A Critical Theory Revisited*. Oxford University Press, 2002.

Floridi, Luciano. “A Look into the Future Impact of ICT on Our Lives.” *The Information Society*, vol. 23, no. 1, Dec. 2007, pp. 59–64. *Taylor and Francis+NEJM*, https://doi.org/10.1080/01972240601059094.

———. “AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models.” *Philosophy & Technology*, vol. 36, no. 15, Mar. 2023, pp. 1–7. *DOI.org (Crossref)*, https://doi.org/10.1007/s13347-023-00621-y.

———. “Data.” *International Encyclopedia of the Social Sciences*, edited by William A. Darity, Macmillan, 2008, pp. 1–13.

———. *Information: A Very Short Introduction*. Oxford University Press, 2010. *academic.oup.com*, https://doi.org/10.1093/actrade/9780199551378.001.0001.

———. "Is Semantic Information Meaningful Data?" *Philosophy and Phenomenological Research*, vol. 70, no. 2, 2005, pp. 351–70. JSTOR.

———. *The Ethics of Information*. Oxford University Press, 2013.

———. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford University Press, 2014.

Földváry Kinga. "Műfaj és adaptáció." *Apertura*, vol. 17, no. 2, 2022, pp. 1–18. *DOI.org (Crossref)*, https://doi.org/10.31176/apertura.2022.17.2.1.

Foucault, Michel. *The Order of Things: An Archaeology of the Human Sciences*. 2nd ed, Taylor & Francis, 2012.

Freise, Matthias. *Inspired by Bakhtin: Dialogic Methods in the Humanities*. Academic Studies Press, 2018. *www.degruyter.com*, https://doi.org/10.1515/9781618117397.

Fuchs, Christian. "Social Media as Participatory Culture." *Social Media: A Critical Introduction*, SAGE Publications Ltd, 2014, pp. 52–68. *DOI.org (Crossref)*, https://doi.org/10.4135/9781446270066.

Gadamer, Hans-Georg. *Igazság és módszer: egy filozófiai hermeneutika vázlata*. Translated by Bonyhai Gábor and Fehér M. István., 2., jav. Kiad, Osiris, 2003.

Genareo, Vincent, et al. "Experiences and Recommendations for Collaborating on Master's of Education Thesis Research." *Literacy Information and Computer Education Journal*, Special 3, no. 1, 2014. *scite.ai*, https://doi.org/10.20533/licej.2040.2589.2014.0239.

Gerbaudo, Paolo. "From Data Analytics to Data Hermeneutics. Online Political Discussions, Digital Methods and the Continuing Relevance of Interpretive Approaches." *Digital Culture & Society*, vol. 2, no. 2, Dec. 2016, pp. 95–112. *DOI.org (Crossref)*, https://doi.org/10.14361/dcs-2016-0207.

Gitelman, Lisa, editor. *"Raw Data" Is an Oxymoron*. The MIT Press, 2013. Infrastructures Series.

Greenblatt, Stephen. "Shakespeare's Leap." *The New York Times*, 12 Sept. 2004. Books. *NYTimes.com*, https://www.nytimes.com/2004/09/12/magazine/shakespeares-leap.html. Accessed 15. Oct. 2025.

———. "The Traces of Shakespeare's Life." *The New Cambridge Companion to Shakespeare*, edited by Margreta De Grazia and Stanley Wells, 2nd ed., Cambridge University Press, 2010, pp. 1–14. *DOI.org (Crossref)*, https://doi.org/10.1017/CCOL9780521886321.001.

Hamacher, Werner. "Afformative, Strike." *Cardozo Law Review*, vol. 13, no. 4, Dec. 1991, pp. 1133–58.

———. "Lingua Amissa: The Messianism of Commodity-Language and Derrida's 'Specters of Marx.'" *Futures*, Stanford University Press, 2002, pp. 130–79. *www.degruyterbrill.com*, https://www.degruyterbrill.com/document/doi/10.1515/9780804779975-007/html.

Han, Byung-Chul. *Infocracy: Digitalization and the Crisis of Democracy*. Translated by Daniel Steuer, Polity Press, 2022.

———. *Psychopolitics: Neoliberalism and New Technologies of Power*. Verso, 2017.

Haslhofer, Bernhard, et al. *Knowledge Graphs in the Libraries and Digital Humanities Domain*. Version 1, 2018. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.1803.03198.

Hassan, Robert. *The Condition of Digitality: A Post-Modern Marxism for the Practice of Digital Life*. University of Westminster Press, 2020. *DOI.org (Crossref)*, https://doi.org/10.16997/book44.

Hayles, N. Katherine. "The Condition of Virtuality." *Language Machines: Technologies of Literary and Cultural Production*, edited by Jeffrey Masten et al., Routledge, 2016, pp.

69–94.

———. “Virtual Bodies and Flickering Signifiers.” *October*, vol. 66, 1993, pp. 69–91, https://doi.org/10.2307/778755. JSTOR.

Heath, Jerome, and Andrej Barovic. “All Prestige Skins in *LoL* and How to Get Them.” *Dot Esports*, 7 Nov. 2024, https://dotesports.com/league-of-legends/news/league-prestige-edition-skins-and-how-to-acquire-them. Accessed 15. Oct. 2025.

Holquist, Michael. “Answering as Authoring: Mikhail Bakhtin’s Trans-Linguistics.” *Critical Inquiry*, vol. 10, no. 2, 1983, pp. 307–19. JSTOR.

Hui, Yuk. “What Is a Digital Object?” *Metaphilosophy*, vol. 43, no. 4, July 2012, pp. 380–95. *DOI.org (Crossref)*, https://doi.org/10.1111/j.1467-9973.2012.01761.x.

Jenkins, Henry. *Convergence Culture: Where Old and New Media Collide*. New York University Press, 2006.

Karcher, Sebastian, et al. “Beyond the Matrix: Repository Services for Qualitative Data.” *IFLA Journal*, vol. 42, no. 4, Dec. 2016, pp. 292–302. *SAGE Journals*, https://doi.org/10.1177/0340035216672870.

Kiséry, András. “Por Se: A Reneszánsz Médiumai.” *Jelenkor: Irodalmi És Művészeti Folyóirat*, vol. 48, no. 11, 2005, pp. 1066–83.

Kitchin, Rob. “Big Data, New Epistemologies and Paradigm Shifts.” *Big Data & Society*, vol. 1, no. 1, Apr. 2014, p. 2053951714528481. *DOI.org (Crossref)*, https://doi.org/10.1177/2053951714528481.

Kittler, Friedrich A. *Discourse Networks 1800/1900*. Stanford University Press, 1990.

———. *Gramophone, Film, Typewriter*. Translated by Geoffrey Winthrop-Young and Michael Wutz, Stanford Univ. Press, 1999. Writing Science.

———. *Operation Valhalla: Writings on War, Weapons, and Media*. Duke University Press, 2021.

———. “There Is No Software.” *The Truth of the Technological World*, Stanford University Press, 2014, pp. 219–29, https://doi.org/10.1515/9780804792622-016.

Krajewski, Markus. *Paper Machines*. Translated by Peter Krapp, The MIT Press, 2011, http://www.jstor.org/stable/j.ctt5hhmbf. JSTOR.

Krämer, Sybille. “From Dissemination to Digitality: How to Reflect on Media.” *Media Theory*, Into the Air, vol. 5, no. 2, 2022, pp. 80–98.

———. “Should We Really ‘Hermeneutise’ the Digital Humanities? A Plea for the Epistemic Productivity of a ‘Cultural Technique of Flattening’ in the Humanities.” *Journal of Cultural Analytics*, vol. 7, no. 4, Jan. 2023, pp. 1–23. *culturalanalytics.org*, https://doi.org/10.22148/001c.55592.

———. “The Cultural Technique of Flattening.” *Metode*, vol. 1, 2023, p. 18.

Krämer, Sybille, and Anthony Enns. *Medium, Messenger, Transmission: An Approach to Media Philosophy*. Amsterdam University Press, 2015. Recursions: Theories of Media, Materiality, and Cultural Techniques.

Kramer, Sybille, and Anita McChesney. “Writing, Notational Iconicity, Calculus: On Writing as a Cultural Technique.” *MLN*, vol. 118, no. 3, Apr. 2003, pp. 518–37. *DOI.org (Crossref)*, https://doi.org/10.1353/mln.2003.0059.

Laskin, Miriam, and Cynthia R. Haller. “Chapter 11. Up the Mountain without a Trail: Helping Students Use Source Networks to Find Their Way.” *Information Literacy: Research and Collaboration across Disciplines*, edited by Barbara J. D’Angelo et al., The WAC Clearinghouse; University Press of Colorado, 2016, pp. 237–56. *DOI.org (Crossref)*, https://doi.org/10.37514/PER-B.2016.0834.2.11.

Leckie, Gloria J., and John E. Buschman, editors. *Information Technology in Librarianship: New Critical Approaches*. 1st ed., www.lu.com, 2008. *DOI.org (Crossref)*, https://doi.org/10.5040/9798400670268.

Leonelli, Sabina. “What Counts as Scientific Data? A Relational Framework.” *Philosophy of Science*, vol. 82, no. 5, Dec. 2015, pp. 810–21. *DOI.org (Crossref)*, https://doi.org/10.1086/684083.

Lotman, Jurij Michajlovič. *Culture and Communication: Signs in Flux: An Anthology of Major and Lesser-Known Works*. Edited by Andreas Schönle, Translated by Benjamin Paloff, Academic Studies Press, 2020. Cultural Syllabus.

Manovich, Lev. *Cultural Analytics*. The MIT Press, 2020.

———. “Database as a Symbolic Form.” *Media Studies*, edited by Sue Thornham et al., Edinburgh University Press, 2009, pp. 236–47. *DOI.org (Crossref)*, https://doi.org/10.1515/9781474473231-027.

———. “New Media from Borges to HTML.” *The New Media Reader*, edited by Noah Wardrip-Fruin et al., The MIT Press, 2003, pp. 13–25.

———. “Seven Arguments about AI Images and Generative Media.” *Artificial Aesthetics: A Critical Guide to AI, Media and Design*, 2023.

McLuhan, Marshall. *Understanding Media: The Extensions of Man*. 1st MIT Press ed, MIT Press, 1994.

Mejias, Ulises A., and Nick Couldry. “Datafication.” *Internet Policy Review*, vol. 8, no. 4, Nov. 2019, pp. 1–10. *DOI.org (Crossref)*, https://doi.org/10.14763/2019.4.1428.

Meyrowitz, Joshua. “Medium Theory.” *Communication Theory Today*, 1994, pp. 50–77.

Minkkinen, Panu. “Media, Cultural Techniques, and the Law: The Other Cornelia Vismann.” *German Law Journal*, vol. 24, no. 9, Dec. 2023, pp. 1597–611. *DOI.org (Crossref)*, https://doi.org/10.1017/glj.2023.115.

Mouffe, Chantal. “Deliberative Democracy or Agonistic Pluralism?” *Social Research*, vol. 66, no. 3, 1999, pp. 745–58.

“New Horizons: Confronting the Digital Turn in the Humanities.” *C2DH | Luxembourg Centre*

for Contemporary and Digital History, https://www.c2dh.uni.lu/projects/new-horizons-confronting-digital-turn-humanities. Accessed 15 Oct. 2025.

Octavia, Lisa. "Living in a Hamster Wheel: Identity Construction through Hopes and Terrors in Bong Joon-Ho's 'Parasite.'" *Rainbow : Journal of Literature, Linguistics and Culture Studies*, vol. 10, no. 1, Apr. 2021, pp. 24–33, https://doi.org/10.15294/rainbow.v10i1.45388.

O'Donnell, Victoria L., et al. "Transition to Postgraduate Study: Practice, Participation and the Widening Participation Agenda." *Active Learning in Higher Education*, vol. 10, no. 1, Mar. 2009, pp. 26–40. *SAGE Journals*, https://doi.org/10.1177/1469787408100193.

Ong, Walter J. *Orality and Literacy: The Technologizing of the Word*. 30th anniversary ed.; 3. ed, Routledge, 2012. New Accents.

———. *Rhetoric, Romance, and Technology: Studies in the Interaction of Expression and Culture*. Cornell University Press, 1971.

Péti Miklós. "'Childe Harold, a rév, mely felé igyekszem' -- Kosztolányi újonnan előkerült Childe Harold fordításáról." *Vigília*, vol. 89, no. 4, 2024, pp. 342–50.

Pfeiffer, K. Ludwig. *A Mediális És Az Imaginárius: Egy Kulturantropológiai Elmélet Dimenziói*. Ráció, 2005.

Picchione, John. "Byung-Chul Han: Digital Technologies, Social Exhaustion, and the Decline of Democracy." *New Explorations*, vol. 3, no. 2, Nov. 2023, pp. 1–23. *DOI.org (Crossref)*, https://doi.org/10.7202/1107743ar.

Pols, Jeannette, et al. "Making Sense with Numbers. Unravelling Ethico-Psychological Subjects in Practices of Self-Quantification." *Sociology of Health & Illness*, vol. 41, no. S1, 2019, pp. 98–115. *Wiley Online Library*, https://doi.org/10.1111/1467-9566.12894.

Pólya Tamás. "A médium mint cselekvési közeg : Médiumelméleti nézőpontváltás a technológiai determinizmus után." *REPLIKA*, no. 76, no. 76, 2011, pp. 51–73.

Primorac, Antonija, et al. “Distant Reading Two Decades On: Reflections on the Digital Turn in the Study of Literature.” *Digital Studies / Le Champ Numérique*, no. Special DSCN Collection #9, Special DSCN Collection #9, Oct. 2023. *www.digitalstudies.org*, https://doi.org/10.16995/dscn.8855.

Romele, Alberto, et al. “Digital Hermeneutics: From Interpreting with Machines to Interpretational Machines.” *AI & SOCIETY*, vol. 35, no. 1, Mar. 2020, pp. 73–86. *Springer Link*, https://doi.org/10.1007/s00146-018-0856-2.

———. *Digital Hermeneutics: Philosophical Investigations in New Media and Technologies*. First issued in paperback, Routledge, 2021. Routledge Studies in Contemporary Philosophy.

———. “The Datafication of the Worldview.” *AI & SOCIETY*, vol. 38, no. 6, Dec. 2023, pp. 2197–206. *DOI.org (Crossref)*, https://doi.org/10.1007/s00146-020-00989-x.

Rosenberg, Daniel. “Data before the Fact.” *“Raw Data” Is an Oxymoron*, edited by Lisa Gitelman, The MIT Press, 2013, pp. 15–40. *DOI.org (Crossref)*, https://doi.org/10.7551/mitpress/9302.003.0003.

Rudin, Cynthia. “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead.” *Nature Machine Intelligence*, vol. 1, no. 5, May 2019, pp. 206–15. *www.nature.com*, https://doi.org/10.1038/s42256-019-0048-x.

Savage, Neil. “Breaking into the Black Box of Artificial Intelligence.” *Nature*, Mar. 2022. *www.nature.com*, https://doi.org/10.1038/d41586-022-00858-1.

Shackell, Cameron. “Finite Cognition and Finite Semiosis: A New Perspective on Semiotics for the Information Age.” *Semiotica*, vol. 2018, no. 222, May 2018, pp. 225–40. *www.degruyter.com*, https://doi.org/10.1515/sem-2018-0020.

Siegert, Bernhard. “Coding as Cultural Technique: On the Emergence of the Digital from

Writing AC." *Grey Room*, vol. 70, Mar. 2018, pp. 6–23. *DOI.org (Crossref)*, https://doi.org/10.1162/GREY_a_00236.

———. *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real*. Translated by Geoffrey Winthrop-Young.

Stiegler, Bernard. *The Fault of Epimetheus*. Stanford Univ. Pr, 1998. Technics and Time / Stiegler, Bernhard 1.

Stiegler, Bernard, and Daniel Ross. *What Makes Life Worth Living: On Pharmacology*. English edition, Polity, 2013.

Taylor, Gary. "Why Did Shakespeare Collaborate?" *Shakespeare Survey: Volume 67: Shakespeare's Collaborative Work*, edited by Peter Holland, vol. 67, Cambridge University Press, 2014, pp. 1–17. Shakespeare Survey. Cambridge Core, *Cambridge University Press*, https://doi.org/10.1017/SSO9781107775572.001.

Tinnell, John. "Grammatization: Bernard Stiegler's Theory of Writing and Technology." *Computers and Composition*, vol. 37, Sept. 2015, pp. 132–46. *DOI.org (Crossref)*, https://doi.org/10.1016/j.compcom.2015.06.011.

Van Der Veer Martens, Betsy. "An Illustrated Introduction to the Infosphere." *Library Trends*, vol. 63, no. 3, Dec. 2015, pp. 317–61. *DOI.org (Crossref)*, https://doi.org/10.1353/lib.2015.0006.

Van Dijck, Jose. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *Surveillance & Society*, vol. 12, no. 2, May 2014, pp. 197–208. *DOI.org (Crossref)*, https://doi.org/10.24908/ss.v12i2.4776.

Van Dijk, Jan. *The Network Society: Social Aspects of New Media*. 2nd ed, Sage Publications, 2006.

Van Dijk, Jan A.G.M. "The One-Dimensional Network Society of Manuel Castells." *New Media & Society*, vol. 1, no. 1, Apr. 1999, pp. 127–38. *DOI.org (Crossref)*,

https://doi.org/10.1177/1461444899001001015.

Viola, Lorella. *The Humanities in the Digital: Beyond Critical Digital Humanities*. Springer Nature, 2023. *library.oapen.org*, https://doi.org/10.1007/978-3-031-16950-2.

Vismann, Cornelia. *Files: Law and Media Technology*. Translated by Geoffrey Winthrop-Young, Stanford Univ. Press, 2008. Meridian Crossing Aesthetics.

Weiland, Steven. "The Scholarly Workflow in the Digital Age: What Do We Know? What Should We Do?" *What's Past Is Prologue: Charleston Conference Proceedings 2017*, 2018. Charleston Conference, *DOI.org (Crossref)*, https://doi.org/10.5703/1288284316687.

West, Sarah Myers. "Data Capitalism: Redefining the Logics of Surveillance and Privacy." *Business & Society*, vol. 58, no. 1, Jan. 2019, pp. 20–41. *SAGE Journals*, https://doi.org/10.1177/0007650317718185.

Winner, Langdon. "Technologies as Forms of Life." *Ethics and Emerging Technologies*, edited by Ronald L. Sandler, Palgrave Macmillan UK, 2014, pp. 48–60. *DOI.org (Crossref)*, https://doi.org/10.1057/9781137349088.

Winthrop-Young, Geoffrey. "Cultural Techniques: Preliminary Remarks." *Theory, Culture & Society*, vol. 30, no. 6, Nov. 2013, pp. 3–19. *SAGE Journals*, https://doi.org/10.1177/0263276413500828.

Wyatt, S. "Technological Determinism Is Dead; Long Live Technological Determinism." *The Handbook of Science & Technology Studies*, edited by E. J. Hackett et al., vol. 7, MIT Press, 2008, pp. 165–81.

Yin, Robert K. *Case Study Research and Applications: Design and Methods*. Sixth edition, SAGE, 2018.

Zhu, Yuxuan. "The Historical Evolution of the Media in McLuhan's Theory." *Cross-Currents: An International Peer-Reviewed Journal on Humanities & Social Sciences*, vol. 8, no.

6, Aug. 2022, pp. 85–90. *DOI.org (Crossref)*, https://doi.org/10.36344/ccijhss.2022.v08i06.002.

Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First edition, PublicAffairs, 2019.

Abstract

In my doctoral dissertation I examine the process and consequences of digitization and datafication from the perspective of culture, knowledge production and epistemology. My central thesis is that the discursive network of digitality in the Kittlerian sense is built on the process of datafication, whose operations can be understood as a specific cultural technique, thus requiring traditional approaches from the humanities, such as hermeneutics, to redefine and reposition in the context of digitality in order to understand the phenomenon.

In the first chapter, titled Theories of Digitality, I provided a general overview of the definitions, theories and approaches to digitalization, with a key role for a general overview of the relationship between technology and culture. Following Bernhard Stiegler, I argue that technology (techné) and knowledge (epistémé) cannot be separated, they always co-create the human subject.

Following that I focused on identifying the characteristics of digitality as a discourse network. After a review of technological theories, this chapter was built on Kittler's media theory to discuss the characteristics of digitality as a discourse network. I identified digital data as the basis of the discourse network, thus necessitating a close examination of the ontology of data, information and knowledge. After that I proceeded to analyze the process of datafication from performative, afformative and hermeneutic perspectives.

Next, in the chapter titled Digital Data as Cultural Technique, my aim was to situate the process of datafication in our cultural, meaning-making system. I argued that the operations of digital data can be approached as cultural techniques specific to the 21st century, rather than as a neutral by-product of technology. The German media theorists Bernhard Siegert, Sybille Krämer, Cornelia Vismann and Markus Krajewksi have identified the logic of the cultural techniques of digitality in historical parallels, and their work can thus provide a background for the study of data-based cultural phenomena in the 21st century.

In the final large section of the dissertation, I used the perspectives of the theoretical framework developed to present four concrete, practical case studies that are an important part of the digital culture of the 21st century, thus becoming cultural techniques of digitality in our everyday practices. Video games, databases, large language models and science fiction literature all represent specific types of digital objects that capture different aspects of digital interactions and data manipulation. The case studies thus highlighted the need for humanities-based approaches to understanding digital culture and digital life.

Összefoglaló

Doktori disszertációmban a digitalizáció és az adatosodás folyamatát és következményeit vizsgálom a kultúra, a tudástermelés és az ismeretelmélet perspektívájából. Központi tézisem, hogy a digitalitás kittleri értelemben vett diskurzushálózata az adatosodás folyamatára épül, amelynek működése egy specifikus kulturális technikaként értelmezhető, ezért a jelenség megértéséhez a humán tudományok hagyományos megközelítéseinek, például a hermeneutikának, újradefiniálására és a digitalitás kontextusában való újrapozícionálására van szükség.

Az első fejezetben, amelynek címe A digitalitás elméletei, általános áttekintést adok a digitalitás definícióiról, elméleteiről és megközelítéseiről, kiemelt szerepet szentelve a technológia és a kultúra közötti kapcsolat általános áttekintésének. Bernard Stiegler nyomán azt állítom, hogy a technológia (techné) és a tudás (episteme) nem választható szét, kulturális értelemben véve mindig együttesen hozzák létre az emberi alanyt.

Ezt követően a digitalitás, mint diskurzushálózat jellemzőinek azonosítására összpontosítok. A technológiai elméletek áttekintése után ez a fejezet Kittler médiaelméletére épül. A diskurzus-hálózat alapjaként a digitális adatot azonosítom, ami szükségessé teszi az adat, információ és tudás ontológiájának alapos vizsgálatát. Ezt követően a performatív, afformatív és hermeneutika perspektíváiból elemzem az adatosodás folyamatát.

A következő, „A digitális adatok mint kulturális technika” című fejezetben célom az volt, hogy az adatizálás folyamatát a kulturális, jelentésalkotó rendszerünkbe helyezzem. Arra a következtetésre jutottam, hogy a digitális adatok működése nem a technológia semleges melléktermékeként, hanem a 21. századra jellemző kulturális technikaként közelíthető meg. A német médiateoretikusok, Bernhard Siegert, Sybille Krämer, Cornelia Vismann és Markus Krajewksi a digitalitás kulturális technikáinak logikáját történeti párhuzamokban azonosították, munkájuk így háttérként szolgálhat a 21. század adat alapú kulturális jelenségeinek tanulmányozásához.

A disszertáció utolsó nagy részében a kidolgozott elméleti keret perspektíváit felhasználva négy konkrét, gyakorlati esettanulmányt mutattam be, amelyek a 21. század digitális kultúrájának fontos részét képezik, és így mindennapi gyakorlatunkban a digitalitás kulturális technikáivá válnak. A videojátékok, adatbázisok, nagy nyelvi modellek és a tudományos-fantasztikus irodalom mind olyan specifikus típusú digitális objektumokat képviselnek, amelyek a digitális interakciók és az adatmanipuláció különböző aspektusait ragadják meg. Az esettanulmányok így rávilágítottak arra, hogy a digitális kultúra és a digitális élet megértéséhez humán tudományokon alapuló megközelítésekre