

3

READING IMAGES: VISUALIZATION AND INTERPRETATION IN DIGITAL TEXT ANALYSIS

NONLINGUISTIC LITERARY ANALYSIS

We typically communicate through language, but often we also communicate with images. How we get information from pictures, however, is different from how we process texts in ways that become crucial when our communication partner is an algorithm. Visualization helps us make sense of the processes of computers and use them to get new information—even and especially when the materials they deal with are not images. This is shown by the field explicitly dedicated to the study of texts: literary analysis. The most recent and innovative research does not use “literary” tools. Instead of reading texts, literary scholars visualize them, utilizing images rather than language.¹

As Franco Moretti and Oleg Sobchuk put it: “If there is one feature that immediately distinguishes the digital humanities (DH) from the ‘other’ humanities, data visualization has to be it.”² DH experts who apply computational techniques to expand the scope and the capabilities of textual analysis are obtaining information with innovative categories and tools—and raising unprecedented issues, such as in studies of the “loudness” of voices in literary texts or of the relationship between the lengths of titles of novels and the size of the market.³ But the results of this literary

analysis are not gained by reading literary texts and are not expressed in literary form. They are not spoken or written—they are shown.⁴

Visualization tools are crucial to the work of the DH, especially in studies involving large amounts of data.⁵ According to Gitelman and Jackson, “Data are mobilized graphically.”⁶ Franco Moretti’s “distant” reading transforms texts into “maps, graphs, trees.”⁷ The patterns identified by algorithms are translated into spatial configurations that transform the complex topology of digital processing into two-dimensional (and possibly three-dimensional) images.⁸ The corresponding techniques are gaining more and more momentum, moving “out of the realm of an exotic research specialty and into the mainstream of user interface application design.”⁹ Münster and Terras propose the phrase “visual digital humanities” as a novel umbrella term in the field.¹⁰

A theoretical analysis of this trend is still not available.¹¹ Why has the textual discipline par excellence, literary analysis, moved toward visual tools?¹² My argument in this chapter is that visualization is an answer to the opacity of algorithmic procedures and a way to make them productive. Instead of explaining (“answering the ‘why’ question”),¹³ the digital humanities are devising other ways to deal with the incomprehensibility of digital processes and to exploit it in interactions with human users. Visualization is a powerful, increasingly widespread, solution.

EXPLORING IMAGES

In itself, there is nothing new in the use of images for communicative purposes. Compared to language (oral or written communication), images have the great advantage of communicating a lot of information at once, if in a less analytic way.¹⁴ As Ware points out, among the greatest benefits of visualization are the sheer quantity of information that can be rapidly interpreted, and the possibility of perceiving emergent properties that were not anticipated.¹⁵ Think of the difference between describing a landscape verbally and presenting a postcard like the image shown in figure 3.1.

A linguistic description takes much longer (one can say only one thing at a time) and includes only the information explicitly taken into account. If you forget to say that there are flowers in front of the house or



Figure 3.1

that the chimney is made of bricks, your interlocutor cannot know it. An image, on the other hand, transmits a great deal of information in a single moment, even information of which neither the sender nor the receiver was aware.¹⁶ Even if the recipient does not actively engage with it, upon seeing the image, they know that the tree is to the left of the house and that there are clouds in the sky. Likewise, visualization can be used to generate for the receiver information that the sender themselves did not know.

With or without a reference text, pictures have always been used in narrations and explanations to carry out two distinct functions: showing information that is already available and making it possible to produce new information. These two functions correspond respectively to the roles of images as *illustrations* and as *visualizations*, which are also central in digital processes.

Illustrations can directly convey information or support a linguistic text (oral or written) to make the communication of information more immediate and persuasive.¹⁷ Illustrating a linguistic text with images, communication takes advantage of both registers: the explicitness of language and the diffusiveness of visual perception.

Images, however, can also be used to autonomously produce information, as a way of “using vision to think.”¹⁸ This is the case specifically

indicated by visualization that I address in this chapter. *Visualizations* are used as a “medium for human interaction with the data”¹⁹—not, as with illustrations, to more efficiently convey already available information, but to create new information. One shows an image, and sees what one gets from it (if one gets something). The image activates a “hypothesis generation process” that could not happen through verbal means.²⁰ Visualization is particularly useful to create knowledge in an interaction with a viewer—not to transmit information but to explore it.²¹ This happens with maps, diagrams, tables, graphs, charts, and with all devices that present visual encodings of data in order to obtain further information.

Visualization is not a tool unique to the humanities. The DH borrows it from the natural sciences, which uses visualization for analytical purposes. It is also an ancient technique to obtain information from data.²² Spatial (bi-, tri-, and today also poly-dimensional) representations allow for the identification of connections and relationships that could not be grasped otherwise. The natural sciences has always done so, exploring and manipulating images, patterns, radiographies, and models.²³ With the intervention of computers, however, the use of visualization has become more complex, and today we distinguish at least three different ways of using images for exploratory purposes. Here I will call them “scientific visualization,” “information visualization,” and “digital visualization.”

Scientific visualization or “scivis” has a long tradition and uses representations to show “physically based” forms.²⁴ Scientific visualizations are bound to the “a priori based spatial layout of the real physical object,”²⁵ which is reproduced in a simplified way through a model, a schema, or a two-dimensional image in order to be explored more easily—think of a map or a geometric drawing. Faithfulness to the world is a requisite. An x-ray plate shows an image of the internal organs to allow diagnosis. A map refers to a territory and reproduces its structure, though not its complexity, and this simplicity is what makes the map so useful for the purpose of orientation. Obviously, the use of images requires skills to interpret them, which may be greatly refined. Even when employing abstract and highly elaborate images, however, scientific visualizations always do so to make visible a structure that exists, but could not be directly perceived.

Information visualization or “infovis” starts from these practices, but differs because it is unbound from the layout of the objects.²⁶ It is a surprisingly recent invention, introduced in the second half of the eighteenth century to enable use of graphics as “instruments for reasoning about quantitative information.”²⁷ The purpose of information visualization is understanding, not representation—gaining “insight, not pictures.”²⁸ Envisioning information has been described as a “cognitive art” that employs abstract, nonrepresentational images to show information, not objects.²⁹ In information visualization, the graphical models may represent concepts and relationships that do not necessarily have a counterpart in the physical world³⁰—time series, frequency of diseases, movements of stock prices, distribution of criminal behavior across generations. The result is the widespread presence of graphs, diagrams, histograms, pie charts, scatterplots, which do not resemble their objects.

The data on which the information is based are quantified and expressed visually with points, circles, rectangles, ascending and descending lines, to allow free exploration and analysis. For example, the evolution of the most common science-fiction themes over time—quite complex and not spatial—can be presented with the lines of a graph.³¹ This makes it possible to see immediately how the popularity of aliens, space travel, robots, and time travel has changed from 1970 to 2009. Geometry and topology are used to express key differences in data with visible signs and with their location in space.³² Exploring the images, the user can develop new ideas.

Since the 1990s the use of information visualization has increased greatly, along with the rise of desktop 2D graphics software and the use of personal computers by designers, and has been further enhanced in the 2000s as a consequence of big data and new high-level programming languages—from which the current DH approaches derive. Analysis focuses on “processes or datasets that are either too large, or too complex, to be fully understood by a single (static) image.”³³ According to Manovich, the use of computers has led to a specific variant of information visualization that can be called *digital visualization*.³⁴ Here a new agent intervenes in the process and allows for endeavors that would not be possible otherwise: this new agent, which explores data and produces

information, is in this case the computer itself. Algorithms do not just show patterns—they find them. Via its digital forms, which involve the autonomous intervention of algorithms in the management and processing of data, visualization changes its object and its purpose.³⁵

The central innovation here is, in my view, that the use of computer-supported visual representations of data is accompanied, explicitly or implicitly, by the promise or the hope to, in Alexandru Telea's words, "discover the unknown."³⁶ Through digital visualization we can obtain knowledge that we were not looking for.³⁷ The autonomous work of algorithms is expected to identify structures (or patterns) in the data without the intervention of the researcher. Visualizing the patterns, then, algorithms can show something the researchers were not searching for, thereby broadening their interpretive horizon.³⁸ What is displayed is not the structure of the objects of a study, nor a simplified representation of the available data, but the configurations autonomously "discovered" by the algorithms, which are offered to interpretation and exploration. The interpretation may then lead to new information.

For literary studies the possibility of using images to produce information opens up new horizons of exploration. Therefore, visualization is taking on a central role in the digital humanities, as today scholars combine the reading of images with the reading of texts.

VISUAL PROVOCATIONS

How does textual analysis change when one uses digital techniques? Why is visualization attaining such a central role in literary studies? The reason lies in the management of incomprehensible materials. DH scholars make extensive use of algorithms to process texts—a process that produces its own texts, though ones that are impenetrable to human readers³⁹—and to work with corpora too large or too small for human analysis. DH programs analyze hundreds of texts or single words and characters within a text.⁴⁰ The unprecedented challenge in managing the results of the working of algorithms is to make informative the outcomes of processes that are often opaque to the human mind.⁴¹

In response to this challenge, experts in literary studies have begun to systematically turn to visualization, which is becoming the fundamental

tool for a new form of textual analysis. This analysis is based on the coordinated contribution of human readers oriented toward meaning, and of algorithmic procedures that do not know and do not use meaning.⁴² Text-processing machines do not think like us and in general do not think at all: “The computer “reads” (processes) the text as a meaningless string of characters.”⁴³ As David Weinberger says, “To imagine thinking the way computers think . . . is to imagine not thinking at all.”⁴⁴ The task of digital visualization is to make these incomprehensible processes informative to human readers; its “critical question” is about the best way to transform data into something that people can understand.⁴⁵

Digital visualization techniques can be seen as exploration tools that allow users to investigate patterns and obtain information—including information that was not there before. Its purpose, for example, can be to “visualize uncertainty” in the patterns that algorithms identify in processed texts.⁴⁶ This is accomplished not by communicating the meanings that patterns and configurations have for the authors of the text (who were not aware of them), nor by communicating the meanings identified in these patterns and configurations by algorithms (for whom texts do not have meaning). Algorithms, after all, certainly do not perceive uncertainty. Instead, visualizations are required because nobody knows what the information generated by algorithmic procedures *is*. Patterns are defined, and information is generated (if it is generated) in interactions with a user who explores the resulting images.

This produces specific challenges. Informative openness is a big advantage of using images, yet at the same time it is a liability the DH have to deal with. Visualization can yield previously unknown information, but one cannot know in advance if and how the visualization will be informative.⁴⁷ One doesn’t know if the user will get information, nor what it will be. Starting from the same data, many alternative views can be produced, which can be more or less informative for the reader.⁴⁸ A visualization designer always has to face the dilemma of “choosing from a multitude of data processing possibilities and an even greater choice of potential visualization options,” which is exacerbated by the fact that the purpose is to identify only the visualizations that can produce “interpretable visual patterns”—that is, those that can be meaningful to the users.⁴⁹

The same data can be illuminating or incomprehensible depending on the technique used. A bubbleline can highlight relationships that are not recognizable in a graph, in a word cloud, or in a histogram, even if the data they visualize do not change.⁵⁰ Behrisch and colleagues show the complexity of choosing from available options to decide how to visualize high-dimensional data.⁵¹ A scatter plot, for example, enables one to see clearly if two variables are correlated, but risks producing visual clutter if large numbers of items need to be displayed. Parallel coordinates and radial visualization, on the other hand, enable analysts to explore patterns across a large set of dimensions, and matrix representations can show patterns at a local and a global level of detail—but a wrong ordering for a specific task may hide the patterns instead of revealing them. Similar considerations apply to all available techniques. Visualization is always an open and problematic process.⁵²

In their use of visualization, DH experts are dealing with these problems. Today there are effective tools to support researchers. For example, the text reading and analysis environment, Voyant, allows users to process corpora of texts by producing many different views: graphs, bubblelines, correlations, mandalas, cirrus, scatterplots, links, DreamScapes, looms, knots, trends, and many others.⁵³ The data underlying the different views are the same, although the resulting images are very different. It is up to the researcher to experiment with the different visualizations and find out what they show—if they show something. Galloway observes that “data have no necessary visual form.”⁵⁴ Visualization is the contingent translation of a mathematical structure into a visual form, and can thus vary in what forms are taken. Even if absolutely controlled, none of them is right or wrong, because “data have no necessary information,”⁵⁵ and “this is information that does not have any obvious spatial mapping.”⁵⁶ Visualization is correct if it works, and this depends on the situation and on the researcher.

Algorithms themselves produce not the results of text analysis, but rather “provocations” that serve as “surprising observations that can challenge existing assumptions.”⁵⁷ The visualizations they show can trigger hypotheses generation,⁵⁸ but the interpretation is up to the scholar dealing with the texts, who can accept the provocation and modify their perspective starting from the “proposals” autonomously generated by

machines—or not. Provocations can work or not work—can generate information or not. If a provocation succeeds, the result is a new form of text analysis, which cannot be attributed only to the researcher but presupposes the active contribution both of the machine processing the materials and of the reader interpreting the results. Using digital visualization, texts can yield information with procedures very different from our familiar reading practices, thus, requiring a reflection on the notion of reading and its forms.⁵⁹

READING, NON-READING, DISTANT READING

The use of algorithms in textual analysis makes it possible to obtain information with methods very different from our established reading practice. Should the application of digital procedures to texts be considered a new way of reading? When one interprets digital images instead of linguistic sentences, is one reading? Who reads, and what?

The answers to these questions depend on what is meant by reading, and in the DH, there is an active debate on the notion of reading and the contribution of algorithms. We are certainly dealing with innovative and potentially very productive methods to manage written materials, which pose a challenge to the established models of literary analysis and criticism.⁶⁰ It is not clear, however, if they are still a form of reading. A leading proponent of expanding our understanding of reading, Franco Moretti, is intentionally ambiguous in this regard. Moretti introduced the very successful term “distant reading” to describe a form of text analysis so different from our familiar practices of reading that it requires “a little pact with the devil: we know how to read texts, now let’s learn how not to read them.”⁶¹ When someone reads at a distance using the visualizations produced by machines, then, do they read or not?

In literary debate the question remains open, with a peculiar notion of reading that “is not ‘really’ reading” and explicitly includes its negation.⁶² In the context of this debate, the ambiguity seems to have a reason. The focus of the discussion lies in the opposition between human *close reading*, dealing with a limited number of texts studied in detail (a “canon”),⁶³ and *distant reading*, as an analysis of units that are “much smaller or much larger than the text: devices, themes, tropes—or genres and systems.”⁶⁴

This requires digital processing of extended corpora that could not be analyzed by a human reader. Consider Franco Moretti's analysis of British novels from 1740 to 1850, which deals with seven thousand titles,⁶⁵ and Lev Manovich's survey on Japanese manga, which works with one million images.⁶⁶ Algorithmic reading is, first of all, distanced from reading the canon, from close reading. Distant reading is non-reading in the sense of not being close up—it's about "zooming out" instead of "zooming in."

Outside of this debate, however, distant reading can also be interpreted as unrelated to reading altogether. Algorithms do not read and do not need to read—this is how they gain their specificity and advantage. Algorithmic text processing is different (distant) from reading on at least two levels: in its relationship with documents and in its management of meaning.

First, the *object* of reading changes. Whereas close reading interprets the text without dissolving its structure, distant reading does the exact opposite.⁶⁷ The traditional notion of reading has a "documentcentric" attitude bound to the unity of the text as a book or an article.⁶⁸ A text as a document is a "communicative event: written by someone, in specific circumstances, to convey a specific meaning. . . . A text is meant to address us, to 'speak' to us." The corpora addressed by distant reading, instead, "are not 'communicative events'"; corpora "do not speak to us"⁶⁹—hence they are not properly documents. If we want to keep corpora as texts, the concept of "text" must be modified, uncoupling it from a restricted reference to individual documents.

The digital humanities are moving in this direction. According to Matthew Kirschenbaum, the materials of algorithmic processing are not documents such as books or articles: "in today's .txtual condition . . . a 'primary record' can no longer be assumed to be coterminous with . . . a 'physical object.'" ⁷⁰ Electronic texts are independent from material supports, that is, from books or newspapers that bind communication to the specific objects circulating the texts. Katherine Hayles points out that in electronic reading, recording media do not coincide with transmission media.⁷¹ The text circulates and is produced independently from the constraints of the book.⁷² Machine-reading devices use materials available on the web to produce their own computational objects, drawing them indifferently from populations of documents, or from individual lines,

nouns or letters.⁷³ This does not mean that a text is produced arbitrarily; rather, it means that its unity no longer depends on the boundaries of the book associated with the author's perspective. It is no longer "document-centric," but depends instead on the constraints of the programs governing the work of algorithms.⁷⁴

This is relevant not so much because we can read a novel in an electronic format on a Kindle or because the novel can include hyperlinks, but because, through data mining and visualization techniques, we can deal with texts that are different from the ones transmitted by books. The stability of printed text is lost in the "processuality" of electronic texts.⁷⁵ What makes a text a text is not the unity of communication related to the intention of the issuer, but its *addressability*⁷⁶—that is, the possibility of being adopted by the machine as a "provisional unity" in its operations. A text is instead what the algorithm processes *as text* at different layers of analysis: analyses of characters, words, lines, works, and of genres.⁷⁷ As Whitmore puts it: "Textuality is massive addressability."⁷⁸

This transformation is connected with a second dimension of "distance" in distant reading: distance *from meaning*. The current way we read has the goal of getting meanings from texts. However, as Moretti says, "Corpora . . . have no meaning in the usual sense of the word."⁷⁹ The meaning of programs is what they do.⁸⁰ Human interpretations referring to meaning provide the starting point for digital processing,⁸¹ but machines do not understand meaning, and their analysis must be independent from the interpretation of each individual researcher⁸²—and strictly speaking from a reference to individuals and their meanings.⁸³ Distant reading moves to a "scopic vision" that is not bound to a single point of view.⁸⁴

Algorithms use meanings as sources of difference that can be combined with one another in a meta-management that does not need to understand the meaning nor the perspective of the author.⁸⁵ Silke Schwandt argues that computers are semantically blind.⁸⁶ Algorithms recognize what a text is about not because they understand its words or interpret the text's meaning, but because they deal with meanings as *things*, identifying formal aspects such as the use of "mine" as an erotic term in Emily Dickinson's texts or the structure of the title of gothic novels.⁸⁷ Meaning is connected to other meanings in order to reveal patterns; but patterns

themselves do not necessarily have a meaning and are not the result of an interpretation.

ALGORITHMIC READING IS NOT ALGORITHMS READING

What, then, does reading mean today, if we want to take into account the contributions of algorithms to this activity? If “machines can read,”⁸⁸ they still read in a different way than human beings, doing so “second hand . . . without a single direct textual reading.”⁸⁹ Algorithms do not do the same things that humans do, but better. Their ability to deal with big corpora is not only a quantitative change: “When we work on 200,000 novels instead of 200, we are not doing the same thing, 1,000 times bigger; we are doing a different thing.”⁹⁰ Dealing with huge corpora, instead of reading, one counts things; instead of interpreting, one builds graphs, maps, and trees.⁹¹ Instead of understanding meaning, one develops a topological analysis that allows pattern visualization “at a distance” that would have escaped the view of traditional close reading. This zoomed out perspective on texts,⁹² which are not themselves read, becomes a “condition of knowledge.”⁹³

Instead of reading texts, DH scholars often observe visualizations—analyzing images rather than interpreting sentences. They could not deal with such materials without the contribution of algorithms. Should we then modify our notion of reading to also include these different things done by algorithms? Like many others, Katherine Hayles thinks that we should.⁹⁴ She argues that we need to expand our understanding of reading and admit a broader repertoire of reading strategies that includes hyperreading as computer-assisted reading, in which linear reading is accompanied by the exploration of links, by search queries, skimming, filtering by keywords and various other electronic management modes.⁹⁵ This understanding should also include *authentic* machine reading, whereby algorithms use digital (possibly unsupervised) methodologies to discover patterns and structures in texts without having had any initial hypothesis.⁹⁶ The option in this case is to expand the notion of reading, assuming a porous boundary between human interpretation and machine pattern recognition.⁹⁷ In this interpretation, reading overlaps

with modeling, gaming, role playing, adapting, translating, rendering, and simulating.⁹⁸

In my opinion, however, this understanding of reading risks becoming so extensive that the notion loses all usefulness. Confronted with the challenge of describing the many complex forms of information processing in our digital societies, we should, instead, hold onto and combine their differences, rather than efface these distinctions in broad notions. The use of algorithmic techniques in the DH prefigures a mode of dealing with texts that does not erase but—on the contrary—accentuates and exploits the differences between different modes of using written materials.⁹⁹ Instead of a porous boundary, we are dealing with a particularly sharp one. According to Katherine Hayles, “Saying computers cannot read is . . . merely species chauvinism.”¹⁰⁰ I prefer the opposite strategy of explicitly claiming that computers do not read and—more crucially—that precisely for this reason, they contribute to reading.

As argued in chapter 1, computers are becoming increasingly effective partners in information processing not because their capabilities resemble ours, but because they are learning to work in more and more distinct ways than humans performing similar tasks. It seems to me that anthropocentric shortsightedness (species chauvinism) occurs today not in denying that machines can be like human beings, but rather in claiming that machines can only be recognized and appreciated for how well they emulate human activities. Human reading does not need to be the standard by which we understand how algorithms process texts. The debate on distant reading shows that they do something different; therefore, combining algorithmic processing and human reading produces a new and powerful way of analyzing texts. Algorithms’ innovative and extremely productive contribution to the production of information relies on their participation in artificial communication.

Instead of reading, Moretti notes, algorithms recognize patterns.¹⁰¹ The difference between algorithmic text processing and reading is highlighted in visualization practices. Algorithms do not read and do not interpret, but instead identify and present patterns to be interpreted. By presenting patterns through visualization, algorithms can make it possible to read otherwise inaccessible texts, such as Gertrude Stein’s “The Making

of Americans,” or to obtain information from corpora that include thousands of texts.¹⁰²

Here the groundbreaking innovation in literary analysis, which marks its difference to conventional reading practices, is in my opinion not simply a dependence on machines and, in general, on non-human devices such as algorithms. With respect to our familiar media, the central innovation is that algorithms are *noisy* media. All other media—whether printed on paper or broadcast over radio waves moving through the air—should be as “silent” as possible, in the sense of transmitting information in a neutral way in which the media themselves are imperceptible. If a medium is perceived in a received communication, as when printed words are not sharp or an image on the screen is blurred, it produces noise—that is, a disturbance that should be minimized.¹⁰³ Digital media can follow this model and practically eliminate transmission noise—for example, in digital music reproduction. But digital technology can also be used differently in communication, making the receiver aware of the active role of the machine and its contribution to the generation of content. The debate over distant reading shows it: “Noise is not an obstacle to interpretation, but its aim.”¹⁰⁴

In distant reading, machine intervention radicalizes McLuhan’s formula of “the medium is the message”:¹⁰⁵ computers are expected to intervene very noisily on content. They should autonomously produce information that differs from that delivered by the participants and which is often completely new. This is a radical innovation, clearly separating digital textual analysis from human forms of reading. While human beings used to be the only ones able to produce information, now digitally supported nonhuman textual analyses produce patterns that can generate new information and enable an unprecedented management of texts.

Nevertheless, algorithms themselves do not read, and reading cannot be accomplished without interpretation. Algorithms only produce patterns, which by themselves are not meaningful, and are generally overabundant. Working with large data sets, such as the corpora on which distant reading is practiced, it is inevitable to find patterns—indeed, to find too many.¹⁰⁶ Algorithms do not need to understand meanings and can work “semantically blind”—“drawing unexpected paths through a documentary space that is distinguished by its overall incomprehensibility.”¹⁰⁷

Without interpretation, however, these incomprehensible patterns are useless.¹⁰⁸ This in my opinion is why visualization, with its different techniques, is becoming central: it permits the use of the “blind reasoning power of computers” to explore patterns and to render them meaningful, and furthermore provides the basis of a new way to analyze text using algorithmic “provocations.”¹⁰⁹

If and when provocations by algorithms are accepted, the resulting textual analysis is a far more complex form of reading. The aim, according to Jessop, is “to support interpretive scholarships by allowing areas or relationships of interest to be identified within large volumes of texts.”¹¹⁰ The interpretation is produced by a human reader, although through ways and potentialities that would not be possible without the autonomous contributions of algorithms. What remains is no longer traditional reading. A scatterplot analyzing the distribution of word forms in a corpus of texts can generate clusters that are not based on interpretation, but can significantly modify interpretive reading—for example, finding connections between words and groups of words in a way that could not be detected by any human observer and thus raising new questions. In these cases, the machine operates as a partner making proposals that can direct interpretation in unexplored directions.¹¹¹

By combining the differing capabilities of human reading and algorithmic processing of texts, one of the most significant methodological innovations of the digital humanities is emerging: an *algorithmic reading* that does not coincide with our traditional interpretive reading and does not imply that algorithms themselves read. It is still a form of reading because it starts from texts and produces interpretations, but in a new, powerful way that relies on the active, autonomous role of algorithms that do not themselves interpret.¹¹² It uses the difference between interpretive reading and algorithmic text processing without opposing or assimilating either.

CONCLUSION

Algorithmic text processing is not in continuity with human meaning-oriented reading.¹¹³ Computers don’t read, they count. Machines don’t understand meaning, they process data. In the DH, literary analysis using

algorithms needs to find a way to make meaningful the results of processes that do not rely on understanding meaning and that are often not in themselves understandable. Instead of trying to interpret them, DH scholars turn to visualization, which can make it possible to obtain from texts information nobody yet knew nor understood—in a manner distinct from both reading and illustration. To analyze written texts, scholars in the DH also observe machine-produced images. The outcome is a new, powerful way of reading texts that relies on practices that are effective precisely because they are not forms of reading. With the contribution of algorithms, digital culture provides us with a form of textual communication that can be enormously informative and even creative—if we accept that the intelligent processes that understand and interpret text are only one component at play in the production of information.

This is a section of [doi:10.7551/mitpress/14189.001.0001](https://doi.org/10.7551/mitpress/14189.001.0001)

Artificial Communication

How Algorithms Produce Social Intelligence

By: Elena Esposito

Citation:

Artificial Communication: How Algorithms Produce Social Intelligence

By: Elena Esposito

DOI: [10.7551/mitpress/14189.001.0001](https://doi.org/10.7551/mitpress/14189.001.0001)

ISBN (electronic): 9780262368865

Publisher: The MIT Press

Published: 2022

The open access edition of this book was made possible by generous funding and support from the MIT Libraries.



The MIT Press

© 2022 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.



Subject to such license, all rights are reserved.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Avenir by Jen Jackowitz. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Esposito, Elena, author.

Title: Artificial communication : how algorithms produce social intelligence / Elena Esposito.

Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series: Strong ideas series | Includes bibliographical references and index.

Identifiers: LCCN 2021013271 | ISBN 9780262046664 (hardcover)

Subjects: LCSH: Telecommunication—Social aspects. | Artificial intelligence—Social aspects. | Online identities. | Social intelligence.

Classification: LCC HM851 .E765 2022 | DDC 303.48/33—dc23

LC record available at <https://lcn.loc.gov/2021013271>

10 9 8 7 6 5 4 3 2 1