The Information Manifold: Why Computers Can't Solve Algorithmic Bias and Fake News by Antonio Badia (review)

Christiana Varda

Information & Culture, Volume 56, Number 2, 2021, pp. 227-228 (Review)

Published by University of Texas Press

The Information Manifold: Why Computers Can't Solve Algorithmic Bias and Fake News

by Antonio Badia

MIT PRESS, 2019. 352 PP.
HARDCOVER, $50
ISBN 978-0-2620-4303-8

REVIEWED BY CHRISTIANA VARDA
CYPRUS UNIVERSITY OF TECHNOLOGY

As we navigate the contemporary digital landscape, information feels ubiquitous. Push notifications from news organizations, social media, and private messaging apps serve as a constant reminder that the flow of information is incessant, if not overwhelming. As we traverse this "information age" characterized by immediate access to abundant information, Antonio Badia invites us to pause and consider what counts as information. His book, *The Information Manifold: Why Computers Can't Solve Algorithmic Bias and Fake News*, examines how we define information by considering three different perspectives (syntactic, semantic, and pragmatic) that color not only how we understand information but also how we approach and manage it online on a daily basis in the context of issues such as algorithmic bias and misinformation.

Badia spends the majority of the book framing his argument that many of the contemporary issues we're facing online can be addressed by understanding how information is viewed across three levels of analysis. At the syntactic level, any data is considered information; the content of the information is of less importance than the transmission of the information from sender to receiver. Building on the ideas of Claude Shannon and Andrey Kolmogorov, Badia underlines that this lens primarily views information as data of any pattern or structure.[1] This view does not consider the message as information, but the semantic level does; from this level, the content and its denoted meaning constitute information. This semantic lens views data as referents and only considers data patterns that are meaningful—as opposed to any pattern at the syntactic level—as information. This semantic approach resembles a more intuitive way of defining information—that the message itself is the information—but it also complicates how information is defined, because meaning is both contextual and tentative. At this semantic level, Badia connects information to knowledge by examining Fred Dretske's view that information is the basis of knowledge and to data by examining Luciano Floridi's view that information is data that we notice and interpret.[2] The semantic approach is somewhat extended by the third, pragmatic level of analysis, which focuses not only on the data patterns that are meaningful but also on the ones that serve a specific purpose. The pragmatic lens is more concerned with the goals of the information, but it still values meaningful patterns. The difference is that it views information as only the data structures or patterns that are relevant to a current situation or content. Badia makes painstaking connections to Shannon's approach to information in the semantic and pragmatic levels and draws a convincing argument on how each of these perspectives works progressively, but more narrowly, in scope.

Beyond these three lenses, Badia discusses the connection between information and communication and, more specifically, the way information is distributed and spread among groups and how it is processed in networks. This network-based approach is an important addition, given that the book goes on to explore issues that are rooted in the social web, and it would have been a serious omission to ignore the communication perspective on information. However, the author argues that despite the value of understanding how information is exchanged and processed, to answer the basic question that this book puts forth, it is important to consider the information content and its relevance to achieving specific goals. One must examine information at the semantic and pragmatic levels before considering how the information is used.

These braided approaches are played out with issues that have recently come to the forefront of the information stage: machine learning, big data, algorithmic fairness, and misinformation. The information problems we encounter today begin with the amount of available data and what we can do with them in order to make better, more "informed" decisions. An attempt to answer these problems is through computing solutions that can help us tackle the abundance of available data, filter them into meaningful patterns, and thus help our decision-making processes. Yet Badia points out that when tackling these issues, computers mostly operate at the syntactic level and view data points as information. The semantic and pragmatic levels are the most important levels for allowing us to find information that satisfies our goals, and for this, human agents are still important. Data form the basis of information, but they need to be interpreted. The human part of the equation in computing solutions allows for meaning in information, but it also bases meaning on assumptions and unwitting biases that are inherent limitations of our information processing behaviors.

Badia concludes the book by problematizing what we consider information as viewed through the three perspectives and underlines the value and limitations of data. He argues that data provide a useful starting point, but they need context; data can lead to different interpretations, which suggests that information is always approximation. This implies that the link between information and knowledge is unreliable, since personal interpretations can color information, even inadvertently. The author further argues that scientific reasoning and knowing how to handle data in an information-saturated context are important skills for negotiating information meaning responsibly. This works as a neat conclusion to the book's thesis, but how one moves from being a passive information consumer to a data-literate citizen is not made clear, even if a good portion of the book is focused on humans' information processing skills and the value of scientific reasoning. Regardless, *The Information Manifold* does manage to pack a lot of multidisciplinary information concisely and provides an excellent reference list that can serve as a starting point for anyone interested in exploring information from a range of disciplinary approaches.

NOTES

1. Though Badia offers no specific citations to Andrey Kolmogorov's work, he engages extensively with works such as Claude Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal* 27, no. 3 (1948): 379–423; Claude Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1963).
2. Fred Dretske, *Knowledge and the Flow of Information* (Cambridge, MA: MIT Press, 1981); Luciano Floridi, *The Philosophy of Information* (Oxford: Oxford University Press, 2011).