Philosophy of Engineering and Technology 8

Hilmi Demir *Editor*

Luciano Floridi's Philosophy of Technology

Critical Reflections

Springer

Luciano Floridi's Philosophy of Technology

Philosophy of Engineering and Technology

VOLUME 8

For further volumes:
http://www.springer.com/series/8657

Hilmi Demir
Editor

Luciano Floridi's Philosophy of Technology

Critical Reflections

Springer

Editor
Hilmi Demir
Philosophy Department
Bilkent University
Bilkent, Ankara, Turkey

ISSN 1879-7202 ISSN 1879-7210 (electronic)
ISBN 978-94-007-4291-8 ISBN 978-94-007-4292-5 (eBook)
DOI 10.1007/978-94-007-4292-5
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2012940529

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The ultimate aim of this volume is to further the philosophical reflection on technology within the context of Luciano Floridi's philosophy of technology. Philosophical reflection on technology is as old as philosophy itself, dating back to the Ancient Greek philosophers. The themes that have dominated the philosophical discourse on technology since then can be roughly categorized into three: (i) the social, cultural, and political impacts of technological developments; (ii) the epistemological status of technological knowledge, especially in relation to scientific knowledge; and (iii) the ontological status of the products of technology, i.e., technological artifacts. Luciano Floridi's philosophy of technology, which is based on his philosophy of information, has something to say about each of these themes. Not only that, his philosophical analysis of new technologies leads to a novel metaphysical framework in which our understanding of the ultimate nature of reality shifts from a materialist one to an informational one, in which all entities, be they natural or artificial, are analyzed as informational entities (Floridi 2010). This is the main rationale behind the choosing of his philosophy of technology as the topic of this volume.

There is no doubt that the information and communication technologies of the twentieth century have had a significant impact on our daily lives. They have brought new opportunities as well as new challenges for human development. According to Floridi, however, this is not the whole story. He claims that these new technologies have led to a revolutionary shift in our understanding of humanity's nature and its role in the universe. By referring to an earlier categorization, he calls this the "fourth revolution." The Copernican revolution was the first, leading to the understanding that we as humans are not at the center of the universe. The second revolution was the Darwinian realization that we are not unnaturally distinct or different from the rest of the animal world. The third was the Freudian revolution, which taught us that we are not as transparent to ourselves as we once thought. With the fourth revolution, says Floridi, "we are now slowly accepting the idea that we might be informational organisms among many agents …, *infORGs* not so dramatically different from clever, engineered artefacts, but sharing with them a global environment that is ultimately made of information, the infosphere. The information revolution [the fourth revolution] is not about extending ourselves, but about re-interpreting who we are" (Floridi 2008a).

This radical claim forms the basis of Floridi's philosophy of technology. Given this basis, philosophical reflection on technology is not only valuable in and of itself, but also brings a completely new framework of analysis for philosophy. In other words, philosophical reflection on technology takes a central role in philosophical analysis. To give an example, Floridi's analysis of object-oriented programming methodology (Floridi 2002), which relies on a method borrowed from a branch of theoretical computer science called Formal Methods, paves the way for defining a new macroethical theory, i.e., Information Ethics. The method he borrows from Formal Methods is the method of levels of abstraction. By using this method, Floridi claims that the moral value of human actions is not different in kind than the moral evaluation of other informational objects. The idea behind the method of levels of abstraction is quite simple and straightforward: the reality can be viewed from different levels. The roots of this simple idea go back to Eddington's work in the early decades of the twentieth century (Eddington 1928). Let me give a brief example in Floridi's own words:

> Suppose, for example, that we interpret p as Mary (p = Mary). Depending on the LoA and the corresponding set of observables, p = Mary can be analyzed as the unique individual person called Mary, as a woman, as a human being, as an animal, as a form of life, as a physical body, and so forth. The higher the LoA, the more impoverished is the set of observables, and the more extended is the scope of the analysis (Floridi 2002).

Perhaps the most crucial feature of the method of levels of abstraction is that the identification relation between two variables (or observables) is never absolute. Rather, the identification is always contextual and the context is a function of the level of abstraction chosen for the required analysis (Floridi and Sanders 2004a).

Floridi utilized his method not only in Information Ethics but also in several other subfields of philosophy. The following quote from his *Minds and Machines* article (2008), in which he responded to some objections raised against the method of levels of abstraction, provides a list of the areas in which the method has been used.

> Jeff Sanders and I were forced to develop the method of abstraction when we encountered the problem of defining the nature of agents (natural, human, and artificial) in Floridi and Sanders (2004b). Since then, we have been applying it to some long-standing philosophical problems in different areas. I have used it in computer ethics, to argue in favour of the minimal intrinsic value of informational objects (Floridi 2003); in epistemology, to prove that the Gettier problem is not solvable (Floridi 2004c); in the philosophy of mind, to show how an agent provided with a mind may know that she has one and hence answer Dretske's question "how do you know you are not a zombie?" (Floridi 2005a); in the philosophy of science, to propose and defend an informational approach to structural realism that reconciles forms of ontological and epistemological structural realism (Floridi 2004b); and in the philosophy of AI, to provide a new model of telepresence (Floridi 2005b). In each case, the method of abstraction has been shown to provide a flexible and fruitful approach (Floridi 2008c).

The jury is still out as to the truth value of the claim stated in the last sentence of this quote. One thing, however, is certain. Floridi's method borrowed from the Formal Methods branch of theoretical computer science and its applications have led to prolific and novel discussions in many different areas of philosophy. For the purposes of this volume, one of the most important applications of the method is in computer ethics. As mentioned above, Floridi claims that his Information

Ethics is a macroethical theory that provides a foundation for computer ethics. His Information Ethics consists of two main theses: (i) information objects *qua* information objects can be moral agents; and (ii) information objects *qua* information objects can have an intrinsic moral value, although possibly quite minimal, and hence they can be moral patients, subject to some equally minimal degree of moral respect (Floridi 2002).

The contributions in Part I of this volume are mainly centered on Floridi's Information Ethics and the method of levels of abstraction. These are Gordana Dodig-Crnkovic's "Floridi's Information Ethics as Macro-Ethics and Info-Computational Agent-Based Models," M.J. Wolf, F.S. Grodzinsky, and K.W. Miller's "Artificial Agents, Cloud Computing, and Quantum Computing: Applying Floridi's Method of Levels of Abstraction," Richard Lucas' "Levels of Abstraction and Morality," and Federica Russo's "The *Homo Poieticus* and the Bridge Between *Physis* and *Techne*."

Dodig-Crnkovic's ultimate aim in her chapter is to provide a general framework for the distribution of moral responsibility in multi-agent systems, which include humans as well as technological artifacts. In order to lay the groundwork for achieving this aim, she starts by providing her own interpretation of Floridi's Information Ethics, which she has been developing since 2006. Her interpretation, called the Info-Computationalist interpretation, is characterized by a recursive self-sustaining loop in which "the bottom-up construction of informational structures gives rise to top-down information re-structuring." In other words, the aggregate of the bottom-level elements forms a collective state that has emergent properties that are not reducible to the properties of the bottom-level informational structures. These emergent properties in turn influence the behavior of all bottom-level structures. Dodig-Crnkovic's interpretation is, to say the least, a novel one, because it allows a structured interaction between different levels of abstraction. In addition to her novel interpretation, she also states the similarities between Floridi's Information Ethics and the pragmatic approach to moral responsibility. The classical analysis of moral responsibility requires an agent with free will, and thus limits the domain of moral responsibility only to humans. In contrast, in the pragmatic approach, moral responsibility is not a result of an individual's duty; rather, it is a role defined by the externalist pragmatist norms of a group. Dodig-Crnkovic claims that Floridi's Information Ethics falls under the category of the pragmatic approach, and in that respect has the potential of providing the foundation for a moral framework in which technological artifacts can be assigned moral responsibility. Armed with these two preliminary explanations, i.e., the Info-Computationalist interpretation and the pragmatic character of Information Ethics, she uses Information Ethics to construct an artificial morality framework in which moral responsibility in intelligent systems is distributed across all agents, including technological artifacts. In her artificial morality framework, moral responsibility is handled as a regulatory mechanism that assures the desirable future behavior of intelligent systems.

Wolf et al.'s chapter, in a sense, is a continuation of an earlier article of theirs that appeared in *Ethics and Information Technology* (2009). In that article, they use two different levels of abstraction for analyzing the ethics of designing artificial agents. Their first level of abstraction, LoA1, is the user's view of an "autonomous system"

such as a software package. The second level is the designer's perception of the system. Their ultimate conclusion in that paper is that the ethical responsibilities of a software designer significantly increase with the development of artificial agents because of the more intricate relationship between LoA1 and LoA2. In their contribution to this volume, they extend their original analysis by introducing a third level of abstraction, LoAS, the level that refers to society's perspective. This is important because new artificial agents not only have effects on individuals but also on the whole society that comprises those individuals. With this new addition, they test the applicability of Floridi's Information Ethics and the method of levels of abstraction to two new computing paradigms: cloud computing and quantum computing. Their overall conclusion is a positive one. They claim that although there are new challenges for Information Ethics in these two computational paradigms, Information Ethics has the potential of successfully meeting those challenges. It should be noted that their chapter also provides a nice and brief overview of the fundamental concepts of quantum computing.

Lucas' chapter is an extensive and detailed criticism of Information Ethics. He criticizes three notions that form the fundamentals of Floridi's theory, which are interactivity, autonomy, and adaptability. Lucas' ultimate conclusion is that Information Ethics, mainly because of being only formally defined, is too artificial and too simple for a natural characterization of morality. Although Floridi thinks that Lucas' understanding of Information Ethics is based on serious misunderstandings and that Lucas' chapter is beyond repair (please see Floridi's reply at the end of this volume), the chapter paves the way for a closer scrutiny of some of the arguments that Floridi has provided in defense of Information Ethics. An example might be helpful at this point. The essential motivation of Information Ethics is to be able to count artificial agents as moral agents. It should be noted that this essential motivation is somewhat different than the motivation behind the earlier characterizations of computer and information ethics. Moor (1985) is a good example of the classic treatment of the subject. In one of their earlier characterizations of Information Ethics, Floridi and Sanders consider a set of possible objections to their main claim about the moral value of artificial agents. These are the teleological objection, the intentional objection, the freedom objection, and the responsibility objection. They then provide counterarguments against those objections. Lucas thinks that none of these counterarguments sufficiently overcome the four possible objections that Floridi and Sanders consider. Of course, whether Lucas is right in his assessment or not is a matter of debate, but Lucas' reasoning urges us to reevaluate the fundamental arguments provided for the philosophical value of Information Ethics. In that respect, it is a valuable contribution to this volume.

Russo, in her chapter, focuses on one particular aspect of Floridi's Information Ethics, the reconciliation of *physis* and *techne* in a constructionist manner. According to Floridi, traditional macroethical theories take the situation which is bound to moral evaluation as given, but this traditional approach ignores the poietic nature of humans as ethical agents. Ignoring the poietic nature of humans is the ultimate basis of the dichotomy between *physis* and *techne* (Floridi and Sanders 2003). The demarcation line between these two has been disappearing because of digital

technologies. Russo agrees with Floridi's analysis and attempts to take the analysis one step further. For Russo, the gradual disappearance of the demarcation line between *physis* and *techne* is not just a result of the new digital technologies; rather, it is dominated by new technologies in general. These new technologies include biotechnology and nanotechnology, which allow us to be "creating altogether new environments that pose new challenges for the understanding of us in the world." Floridi's Information Ethics, according to Russo, successfully accounts for the ethical implications of these new technologies, but, she continues, the epistemological implications are also at least equally important and need to be analyzed. This is what she aims to achieve in her chapter. In that respect, it would not be wrong to say that Russo takes Floridi's original analysis of digital technologies and applies it to a wider domain.

The two chapters in Part II provide novel ways of categorizing scientific and technological advancements on the basis of metrics different than Floridi's metric, which is based on introverted effects of scientific changes on the way we understand human nature. These are Anthony F. Beavers' "In the Beginning Was the Word and Then Four Revolutions in the History of Information" and Valeria Giardino's "I Mean It! (And I Cannot Help It): Cognition and (Semantic) Information."

Beavers, in his chapter, gives us a different categorization of the technological revolutions that mankind has experienced in its entire history. As mentioned above, Floridi's categorization of the information revolution as the fourth revolution is based on the metric of the way scientific developments change our understanding of ourselves. Thus, according to this metric, scientific developments that have led to a reassessment of humanity's fundamental nature and role in the universe are counted as revolutionary. Of course, as Floridi himself states, other metrics are also possible. In his chapter, Beavers offers a different metric that is not supposed to be an alternative to Floridi's metric, but rather complementary. The suggested metric is the history of information flow itself. In other words, technological and scientific advancements are categorized according to "the kind of information that can be stored and transmitted, the speed of information transmission, its preservation, and its reach." This metric also gives us four revolutions: the Epigraphic Revolution, the Printing Revolution, the Multimedia Revolution, and the Digital Revolution. The last one, which corresponds to Floridi's fourth revolution, is characterized by the introduction of automated information processing. There are two interesting features of Beavers' categorization that I would like to mention in this short preface. The first is that in his categorization, the Digital Revolution is not considered as a discontinuity from the previous revolutions, because information transmission and coding were also present, albeit in different forms, in the previous revolutions. What the Digital Revolution has brought to the table is new and revolutionary technological affordances that are made possible by automated information processing. This interesting feature, perhaps, is what fundamentally differentiates Beavers' categorization from Floridi's categorization. The second point is that the trajectory of the history of information flow is not characterized merely by the evolution of particular technologies, but also by the evolution of the informational networks that those particular technologies enable. After establishing his new

categorization, Beavers situates the role of Philosophy, in particular the role of Philosophy of Information, in the historical context of the categorization by providing both valuable historical insights for the evolution of philosophical analysis and crucial questions that will help in the advancement of the Philosophy of Information as a new *philosophia prima*.

Giardino, in her chapter, also provides a different categorization of technological revolutions. Giardino argues that what Floridi calls the fourth revolution is in fact the second information revolution. The underlying reason for this difference is her analysis of information from a cognitive perspective. She thinks that we have been living in an informational environment all along, and that the infosphere includes all informational cognitive agents and cognitive tools. Given this understanding, any artifact that aids symbolic activities becomes an informational cognitive tool. Humans have been living in an informational environment since the time of the invention of the first tool that aided symbolic thinking. For Giardino, the correct characterization of information revolution(s) should be based on how information is transmitted across generations. Thus, the first information revolution is characterized by the transformation of the transmission from sequences of DNA to cultural transmission. The second information revolution, i.e., Floridi's fourth revolution, is characterized by the switch from cultural transmission to online transmission, according to Giardino. One of the valuable features of this chapter, among many others, is its interdisciplinary character. Giardino nicely brings together the literature on Philosophy of Information with the literature on Developmental Psychology and the literature on Cognitive Science.

The contributions in Part III take Floridi's Philosophy of Technology and Philosophy of Information as their basis and apply them to different domains: Elena Pasquinelli's "What Happens to Infoteachers and Infostudents After the Information Turn?" to education, Raphael Cohen-Almagor's "Content Net Neutrality: A Critic" to the regulation of freedom of speech on the Internet, and Armando Malheiro da Silva and Fernanda Ribeiro's "Information Science and Philosophy of Information: Approaches and Differences" to Information Science.

With the changes brought about by the information revolution, we humans have become inforgs that live in the infosphere, according to Floridi. The information revolution has led to a reontologization of our ordinary environment, where the divide between online and off-line has been disappearing. Our environment, the infosphere, "will become increasingly synchronized (time), delocalised (space) and correlated (interactions)," says Floridi. Pasquinelli, in her chapter, in light of Floridi's description of the infosphere, analyzes the past and possible future effects of the information revolution on educational institutions, practices, and actors. She starts her chapter with a diagnosis: the information revolution has not yet revolutionized education. For her, the main reason for this is the reluctance of educational institutions and actors in adopting the new tools, approaches, and paradigms that are made possible by information and computational technologies, especially in comparison to the institutions and actors of other domains. She then compares two different ways of changing the educational institutions and practices. The first is the top-down approach mostly adopted by policy makers. She cites the "One Laptop per Child"

(OLPC) program as an example of the top-down approach and shows the difficulties involved in changing educational practices in this way. According to Pasquinelli, change from the top is difficult, mainly because of the sheer size of educational institutions and the long tradition of educational paradigms and practices. A second reason, which is clearly seen in the OLPC case, is that top-down changes usually do not include students, who are the ultimate users of education, in the design of changing programs. Then she proceeds to give an example of a bottom-up approach that she claims to be more promising. Her fascinating example is the experience of Math on MXit from South Africa. With this example, she urges educational institutions and actors to implement the new technologies from the bottom up. The ultimate goal of such changes, for her, is to turn students into infostudents and teachers into infoteachers. During this transformation, which will be slow and gradual, she says, the old paradigms of education will be challenged because of the new tools and approaches of the information revolution. As the dominant example of the old educational paradigms, she gives the Victorian school, which was defined by the following three characteristics: (i) a dedicated and separated space for learning, (ii) a dedicated time for learning, and (iii) well-defined roles for the learner and the teacher. With the Internet, mobile phones, and digital media, she says, learning could occur anywhere and anytime. Moreover, the demarcation line between the student and the teacher will be blurred to the point of disappearance. In short, Pasquinelli's chapter is an informative and fascinating one in which she urges us to reontologize and reconceptualize our environment for education.

Cohen-Almagor, in his chapter, uses Floridi's Information Ethics in order to identify the democratic regulative principles of freedom of speech on the Internet and the responsibility of Internet Service Providers and Web Hosting Services. He starts his analysis by distinguishing three different senses of "net neutrality": (i) net neutrality as a nonexclusionary business practice; (ii) net neutrality as an engineering principle, allowing traffic on the Internet in a nondiscriminatory manner; and (iii) net neutrality as content nondiscrimination. He calls the third sense Content Net Neutrality. Although he accepts the first two senses as the fundamental principles that should underlie Internet regulation, he rejects Content Net Neutrality. Following Floridi's proactive approach to Ethics, which states that the ethical obligation in the information age is not limited to ethical behaviors in the infosphere but needs to extend to actively shaping the infosphere for the betterment of the humanity, Cohen-Almagor urges us to regulate the available content on the Internet. He argues that content that is morally repugnant and/or at odds with democratic ideals should not be made available on the Internet, and that the primary responsibility for this lies with Internet Service Providers and Web Hosting Services. Throughout his discussion, he uses several striking examples that seem to support his position.

As Silva and Ribeiro point out, Information Science as an autonomous field of study that appeared in the late 1950s. Since then, this new field of inquiry, which could be seen as a continuation of the library sciences, has seen an immense and rapid growth. Despite this rapid growth, however, its nature has not yet been precisely defined. This is perhaps due to the inherently interdisciplinary character of the field. Most interdisciplinary fields, for example Cognitive Science, have gone

through a similar stage of development. Silva and Ribeiro, in their chapter, provide an all-encompassing framework for the nature and identity of Information Science. In their framework, Information Science is "a unitary yet transdisciplinary field of knowledge, included in the overarching area of the human and social sciences, which gives theoretical support to some applied disciplines such as Librarianship, Archivistics, Documentation and some aspects of Technological Information Systems." After providing their framework, they turn to Floridi's Philosophy of Information with the aim of finding a firm philosophical grounding for Information Science. While doing that, they state their own definition of information, which implies the following properties: structured by an action, integrated dynamical, has potentiality, quantifiable, reproducible, and transmissible. Their definition of information has some differences from Floridi's definition of semantic information. Perhaps one of the crucial differences is their distinction between informational data and noninformational data. The analysis of the differences and similarities between their definition of information and Floridi's semantic information is by itself valuable. Moreover, along the way they also bring together different threads of discussions, ranging from the French philosopher Ruyer's work on visual sensation to Søren Brier's Cybersemiotics. Given their analysis of Information Science and the connections they identify between Information Science and Philosophy of Information, it is plausible to conclude that Information Science could be understood as applied Philosophy of Information.

The main focus in Part IV is the epistemic and ontic aspects of Floridi's Philosophy of Information. The contributions here are Eric T. Kerr and Duncan Pritchard's "Skepticism and Information," Joseph E. Brenner's "Levels of Abstraction; Levels of Reality," and Steve T. McKinlay's "The Floridian Notion of the Information Object."

It is almost a truism to say that information should be "adequately created, processed, managed and used" (Floridi 2010). The bombardment of information that we all face in this day and age requires proper information management. As rightly pointed out by Kerr and Pritchard, proper information management requires paying attention to the connection between information and knowledge. After all, information is valuable as long as it paves the way for the acquiring of knowledge. In their chapter, Kerr and Pritchard focus on this important issue, i.e., the epistemic value of information. One of the milestones in the literature on the epistemic value of information is Dretske's book *Knowledge and the Flow Information*, in which a comprehensive epistemology based on information is provided. One of the controversial features of Dretske's framework is its denial of the principle of epistemic closure, which simply states that if an agent knows a proposition and knows that the proposition in question implies another one, then the agent also knows the implied proposition. Dretske's main reason behind the denial of closure is that, for him, information about appearances can never completely rule out skeptical doubts. Kerr and Pritchard claim that Dretske is wrong and that there are ways in which information could address skeptical doubts. They examine two such ways in their chapter: Ram Neta's contextual approach and John McDowell's disjunctivism. Kerr and Pritchard's chapter is valuable in and of itself. Moreover, it opens doors for a different

approach to the epistemic value of information. Dretske's epistemological analysis is done in a hybrid context of doxastic and informational concepts. Kerr and Pritchard's analysis of the closure principle may also be understood as showing a need for moving to a purely informational context of analysis for knowledge, and this is exactly what Floridi does in his Philosophy of Information.

In his chapter, Brenner provides an extensive comparison of his logico-ontological theory, which is called Logic in Reality, and Floridi's Philosophy of Information. According to Brenner, "the broad theory of information proposed by Floridi requires an understanding of the properties and role of information at all levels of reality, in all entities." In other words, a complete theory of information should clarify the relevant ontological properties of information. Given the Kantian spirit of his theory, however, Floridi is quite cautious in making any ontological commitment about reality and entities. The method of levels of abstraction is proposed as a more inter-subjective, socially constructible (hence possibly conventional), dynamic, and flexible way to further Kant's approach. This method, claims Floridi, needs to be seen as a step away from internal realism, but this does not imply that it is a step toward external realism (Floridi 2008b). Thus, according to Brenner, in its current status Floridi's Philosophy of Information seems to be incomplete. Brenner claims that his Logic in Reality remedies this problem and complements Floridi's theory, and he discusses this at length in his chapter. To put it briefly, Logic in Reality is an extension of logic to complex real processes, providing a framework for analyzing and making inferences about complex real world entities and processes at all levels of reality, including biological, cognitive, and social levels. It is obvious from this nutshell definition that the processes that Logic in Reality aims to address include information production and transfer, as well. Some of the philosophically interesting features of Brenner's Logic in Reality are as follows. First, the proposed logic is nonpropositional and non-truth-functional. Second, it is grounded in a fundamental dualism, dynamic opposition, that is claimed to be inherent in energy and present in all real phenomena. In other words, real complex phenomena are in a contradictional relation between themselves and with their opposites. Third, the dynamic opposition in energy is accompanied by the law of the included middle, and thus there is no room for the principle of noncontradiction. Fourth, Logic in Reality neither requires nor commits to abstract categorical structures that separate different aspects of reality. Thus, most of the absolute distinctions of the traditional philosophical analysis, such as the one between epistemology and ontology, disappear in the framework of Logic in Reality. Fifth, Logic in Reality is based on a process-ontological view of reality, which means that the ontological inventory of the world is composed of processes at different levels of complex real phenomena. A direct result of this fifth feature is that Brenner's Logic in Reality implies an ontological levelism. As clearly stated in his defense of Informational Structural Realism (Floridi 2008b), Floridi is committed to the epistemological levelism that his method of levels of abstraction implies, but, as a result of his Kantian general framework, he finds ontological levelism untenable (Floridi 2008c). Brenner states that the ontological levelism that Floridi finds untenable is a result of the misconception of reality as seen through the glasses of classical logic and the traditional object-based

ontological approach. In other words, according to Brenner, any ontological levelism that is based on an absolute distinction between epistemology and ontology is untenable, as Floridi rightly argues, but once the epistemology/ontology of Logic in Reality is adopted, then the ontological levelism becomes tenable and compatible with Floridi's Philosophy of Information.

As the ontological basis of his Philosophy of Technology, Floridi defends a form of structural realism which he calls Informational Structural Realism. In this particular version of structural realism, objects are considered as structural entities which are nothing but a collection of data clusters. This gives rise to Floridi's notion of informational objects as the fundamental ontological entities. As a side remark, it should be noted that although Floridi's analysis of objects in informational terms is quite novel, the history of including information as a fundamental entity in the metaphysics of the world dates back to Wiener's work on *Cybernetics* (Wiener 1948). In order to establish his notion of informational object, Floridi heavily relies on the lessons that he draws from object-oriented programming (OOP), both in terms of the methodology and of the ontology of OOP in constructing his Philosophy of Information and therefore his Philosophy of Technology. McKinlay's chapter focuses on the similarities and differences between OOP and Floridi's Philosophy of Information with respect to their ontology. McKinlay claims that the objects of OOP cannot be the informational objects that Floridi needs in his ontology simply because the objects of OOP are referents, whereas Floridi's informational objects are supposed to be ontologically primitive. McKinlay's claim is almost a direct result of his nominalism about conceptual objects such as OOP classes. His defense of nominalism heavily draws upon Quine's ideas. In addition to its philosophical value in terms of calling our attention to the ontological issues surrounding information and artifacts, McKinlay's chapter also provides a nice introduction to object-oriented programming.

The last chapter of the volume, "The Road to the Philosophy of Information," is Floridi's reply chapter in which each of the contributions are critically evaluated.

This volume, in my humble opinion, is quite promising in terms of achieving its ultimate aim, which is to further the philosophical reflection on technology.

References

Eddington, Arthur. 1928. *The nature of the physical world*. Cambridge: Cambridge University Press.

Floridi, Luciano. 2002. On the intrinsic value of information objects and the infosphere. *Ethics and Information Technology* 4: 287–304.

Floridi, Luciano. 2003. On the intrinsic value of information objects and the infosphere. *Ethics and Information Technology* 4(4): 287–304.

Floridi, Luciano. 2004b. The informational approach to structural realism. final draft available as IEG – Research Report 22.11.04. http://www.wolfson.ox.ac.uk/~floridi/pdf/latmoa.pdf

Floridi, Luciano. 2004c. On the logical unsolvability of the gettier problem. *Synthese* 142(1): 61–79.

Floridi, Luciano, and Jeffry W. Sanders. 2004a. On the morality of artificial agents. *Minds and Machines* 14(3): 349–379.

Floridi, Luciano, and Jeffry W. Sanders. 2004b. On the morality of artificial agents. *Minds and Machines* 14(3): 349–379.
Floridi, Luciano. 2005a. Consciousness, agents and the knowledge game. *Minds and Machines*. 15(3–4): 415–444.
Floridi, Luciano. 2005b. Presence: From epistemic failure to successful observability. *Presence: teleoperators and virtual environments* 14(6): 656–667.
Floridi, Luciano. 2008a. Artificial intelligence's new frontier: Artificial companions and the fourth revolution. *Metaphilosophy* 39(4/5): 652–654.
Floridi, Luciano. 2008b. A defence of informational structural realism. *Synthese* 161(2): 219–253.
Floridi, Luciano. 2008c. The method of levels of abstraction. *Minds and Machines* 18: 303–329.
Floridi, Luciano. 2010. *The philosophy of information*. Oxford: Oxford University Press.
Floridi, Luciano, and Jeffry W. Sanders. 2003. Internet ethics: The constructionist values of *Homo poieticus*. In *The impact of the internet on our moral lives,* ed. R. Cavalier, 195–214. New York: SUNY.
Moor, James. 1985. What is computer ethics? *Metaphilosophy* 16(4): 266–275.
Wiener, Norbert. 1948. *Cybernetics: or control and communication in the animal and the machine*. New York: Technology Press/Wiley.
Wolf, M.J., K. Miller, and F.S. Grodzinsky. 2009. On the meaning of free software. *Ethics and Information Technology* 11(4): 279–286.

Contents

Part I
Information Ethics and the Method of Levels of Abstraction

Chapter 1
Floridi's Information Ethics as Macro-ethics and Info-computational Agent-Based Models

Gordana Dodig-Crnkovic

1.1 Introduction

> *There are, however, "correct accounts" that may complement and reinforce each other, like stones in an arch.* Floridi (2008a, b, c, d)

Ten years after the introduction of Information Ethics (IE) which is an integral part the Philosophy of Information (PI) (Floridi 1999, 2002), Floridi's contribution to the subsequent production of knowledge in several research fields has been reviewed. Among others, two recent special journal issues dedicated to Floridi's work, *Ethics and Information Technology*, Vol. 10, No. 2–3, 2008 edited by Charles Ess and *Metaphilosophy*, Vol. 41, No. 3, 2010 edited by Patrick Allo witness the vitality of his research program of PI. It is far from a closed chapter in the history of philosophy. Contrariwise, it is of great interest for many researchers today, and its development can be expected to contribute to the elucidation of a number of central issues introduced or enhanced by Information and Communication Technologies, ICT.

For IE, moral action is an information processing pattern. It focuses on the fundamentally informational character of reality (Floridi 2008a) and our interactions with it. According to Floridi, ICTs create our new informational habitat "*constituted by all informational entities (such as informational agents, their properties, interactions, processes and mutual relations)*" which is an abstract equivalent of an eco-system. IE is thus a generalization of environmental ethics towards a:

– less anthropocentric concept of agent, including non-human (artificial) and distributed (networked) entities

G. Dodig-Crnkovic (✉)
School of Innovation, Design and Engineering, Computer Science Laboratory,
Mälardalen University, Västerås, Sweden
e-mail: gordana.dodig-crnkovic@mdh.se

H. Demir (ed.), *Luciano Floridi's Philosophy of Technology: Critical Reflections*,
Philosophy of Engineering and Technology 8, DOI 10.1007/978-94-007-4292-5_1,

– less biologically biased concept of patient as a 'centre of ethical worth' in any form of existence.
– more inclusive conception of environment that encompasses both natural and artificial eco-systems.

As moral judgments vitally depend on the information about what the case is and what is understood to be a desirable state of affairs, the macro-ethical behavior of networks of agents depends on mechanisms of information processing and communication. Moral responsibility increases for an agent who gets better informed. Information streams in the Infosphere can both enrich and pollute the informational environment for an agent. Those informational processes are essential in the analysis of behaviors of networks of agents, biological and artificial.

Classical ethics approaches typically look at individual (e.g. Virtue Ethics) or group behavior (e.g. the Ethics of Rights) while IE gives a framework for an agent-based approach. It is important to notice that Floridi's Philosophy of Information with Information Ethics is a research program and not a single theory. As a macro-ethics, applicable to networks of communicating agents and at the same time giving a fundamental-level view of information patterns and processes, IE can help identify general mechanisms and understand their workings. The insight into the underlying informational machinery helps to improve our analysis of ICT-enhanced systems. It is now possible to study the effects of different types of information communication, and their influence on informational networks, including the role of misinformation, disinformation, censorship of information (lack of information) and similar.

1.2 Info-computationalist Perspective on Some Basic Ideas of Information Ethics

In what follows, I will present examples of agent-based analysis of IE in socio-technological systems, elucidating ethical issues of IE within the Info-Computationalist framework as defined in Dodig-Crnkovic (2006a, 2009, 2010) and Dodig-Crnkovic and Müller (2010). That will say I will try to emphasize the diversity of existing ethical approaches, their mutual relations and the role IE plays in a deeper understanding of ethical conditions, based on dual-aspect ontology with information as a structure and computation as a process. In this reading, the contribution of IE is primarily within meta-ethics, but it sheds new light even on normative and descriptive ethics as well as on applied ethics.

IE provides a conceptual space and analytic tools for addressing the dynamic/cybernetics character of relationships between information objects. This approach helps establishing links between information, knowledge and practices of ethics. The proposed Info-Computationalist interpretation reveals a recursive self-sustaining loop: bottom-up construction of informational structures giving rise to top-down information re-structuring (emergent property). Bottom level information elements

produce – through mutual interactions – a collective state that in its turn influences the behavior of each of the bottom-state elements. It should be emphasized that this mechanism, though exhibiting circularity, does not produce "vicious circles" as it stands in a continuous interaction with the environment which provides variation.[1]

The explication of the role of IE is based on the following Info-Computational elements:

1. Ontology is informational; the fabric of reality is (proto) information. (Informational Structural Realism, (Floridi 2008a))
2. Being is process of (natural) computation = Being is information processing, based on natural computing, which is both digital and analog. (Pancomputationalism 2009)
3. Information (structure) and computation (process) are two basic complementary concepts that constitute dual-aspect ontology.
4. Informational structures are physical; there is no information without physical implementation.
5. Based on physical laws, informational structures interact, evolve, and build more and more complex constellations, especially in intelligent living organisms that use "raw information"/(proto) information from the world to construct knowledge and form decisions. (Info-Computational Naturalized Epistemology (Dodig-Crnkovic 2008))
6. Ethical norms are among mechanisms that humans have developed in order to provide guidance in decision making and conduct. They can be understood as a result of successive evolution of info-computational structures in goal-driven living organisms.
7. Informational structures constitute complex systems which can be analyzed on different levels of organization/levels of description/levels of abstraction. IE is the first ethical approach focused on the fundamental level of information.

The above is based on the following fundamental principles, defined in Dodig-Crnkovic and Müller (2010)

(IC1) The ontologically fundamental entities of the physical reality are information (structure) and computation (change).
(IC2) Properties of a complex physical system cannot be derived solely from the properties of its components. Emergent properties must be taken into account.
(IC3) Change of informational structures is governed by laws.
(IC4) The observer is a part of the system observed.

[1] Among physical systems, living organisms are known to use this type of mechanisms in diverse contexts, such as metabolism, reproduction, growth end alike. On a theoretical level, Computing with Computer Science as its subset presents a rich source of examples of self-referential, circular systems that are not vicious, but perform intelligible functions like e.g. program loops, fractals and other recursive functions.

The idea of Levels of Abstraction is central to PI and even to IE, so in what follows I will try to frame the Info-Computational reading of the role of LoA in PI.

1.2.1 On the Concept of Levels of Abstraction

One of the most important insights of PI and IE is their explicit addressing of different Levels of Abstraction/Levels of Organization/Levels of Description of analysis:

> *LoAs are teleological, or goal-oriented. Thus, when observing a building, which LoA one should adopt -architectural, emotional, financial, historical, legal, and so forth – depends on the goal of the analysis. There is no "right" LoA independently of the purpose for which it is adopted, in the same sense in which there is no right tool independently of the job that needs to be done.* (Floridi 2008a, b, c, d)

Epistemologically LoA depends on the *type of interaction* between the cognizing agent and the object of the study. The type of interaction is in its turn defined by the teleological nature of knowledge production/acquisition.

Historically, research fields have typically addressed one level of abstraction/organization/description of reality. There are microscopes and there are telescopes, and visible with the help of those instruments/research tools are their specific worlds. In a microscope, no stars are visible, and in a telescope, no atomic structures. Why is it not common for a framework to encompass several levels of description e.g. to start with a very basic level of organization and encompass all levels up to macroscopic ones? For each of the layers, emergent properties show up as a result of systemic organizational phenomena. The difference between information and knowledge is not the difference in stuff but the difference in organization (structure). Looking into knowledge with fine resolution, one will only find information. Likewise, looking at the world through informational spectacles one will only see information in different constellations. Looking at a human with fine resolution, one will find only atoms, which again are known to us as information.

Every level of organization/level of complexity/level of abstraction has its own "rules of the game" and every new one emerges from the previous ones. The classical ethical discourse uses conceptual repertoire based in everyday human experience. The following passage from Hongladarom (2008) addresses the movement from the level with maximum abstraction of PI towards the detail-rich world of everyday life in the analysis of individual's right to (informational) privacy.

> *And here we are descending from the level of abstraction toward the greater specificity of everyday reality. Even if we believe that ontology is constituted by information, since reality can be described in more and more details and at deeper levels of abstraction, thus necessitating the need for more information, the need to protect privacy would not be affected because there being the Infosphere as basic reality does not mean that all information should be in the hands of the political authority. The question about Infosphere and privacy is designed to illustrate a challenge of the anti-naturalist who emphasizes the putative possibility of the individual against the ontology, but the two need not be in conflict with each other.*

Some critics feel uneasy with the Levels of Abstraction in fear of ethical relativism, but the fear is unfounded. Defining the Level of Abstraction adds to our understanding

of a model. An analogy with natural sciences is instructive. Physics has specific models of the world on many different Levels of Abstraction: from elementary particles, atoms, molecules, solid state, classical mechanics and fluid dynamics, astrophysics to cosmological level. There is also a remarkable emerging field of complex systems which is not only about phenomena on specific levels of organization, but also deals with interactions among different levels. As a result, a complex system as a whole exhibits properties that are distinct from the properties of its individual parts. PI uncovers similar complex structures in epistemology and ontology while IE does the same for ethics. This makes IE a promising research program, and its practical applications are already many and will surely increase in number and importance.

1.2.2 On the Idea of Good in Information Ethics

One of frequent misunderstandings of IE is related to the intrinsic value of informational objects, which in its turn is connected to the understanding of the Levels of Abstraction of a model. A common misconception that follows this confusion is that IE will provide machinery for automatization of ethical decision-making. However, being on a fundamental level, IE will in the first place help us understand basic structures and underlying mechanisms. IE in relation to traditional ethical approaches is like molecular biology in relation to classical biology. We do not expect molecular biology to give us all answers on questions of the living world, but it provides a solid underpinning for the rest of biology. As in other research fields, the diversity of ethical approaches is still equally valuable, and it presupposes human judgment and interaction among theoretical structures.

Informational objects are a priori valuable. *If nothing else is known, we are advised not to destroy or distort informational structures*. On the higher levels of organization, such as the human one, it might well be that we must clean our mail inboxes or hard discs and that is of course not ethically problematic. Respect for information is grounded in respect for nature. One should not destroy natural objects without good reason. Nonetheless, that does not imply that we are not allowed to change anything in the world.

Hongladarom (2008) finds parallels of Floridi's IE with Spinoza's ethics in their ethical naturalism, and concludes that variety of approaches is after all inevitable. On the level of everyday practices, unity in diversity is naturally achieved through interactions:

> *What this translates to the contemporary situation of information ethics is that there are always bound to be many different ways of conceptualizing one and the same reality, and it is the people's needs, goals and desires that often dictate how the conceptualizing is done. However, when different groups of people interact, these systems become calibrated with one another. This is possible because they already belong to the same reality.*

Among the criticisms of IE, Capurro (2008) focus on the intrinsic value of informational objects, Brey (2008) makes a proposal to modify IE from a value-based into a respect-based theory in order to agree with the received view that

"*inanimate things in the world deserve moral respect, not because of intrinsic value, but because of their (potential) extrinsic, instrumental or emotional value for persons*", while Søraker (2007) proposes attribution of relational value to informational objects making the distinction between intrinsic, relational, and instrumental value. All critique points towards humans as a nexus of our ethical interest, which PI is from the outset constructed to avoid:

> *IE adopts this informational ontology (or better: the corresponding LoA) as a minimal common denominator that unifies all entities.* (Floridi 2008a, b, c, d)

This move towards connecting PI's decentralized universal perspective with classical ethical and human-centered approaches is, however, justified and necessary. We as a civilization are (still) "only" humans and our way of cognizing the world is (still) "only" human, so even if we at times adopt a fundamental level of informational structures and processes, it is in the first place in an attempt to understand the basic underlying mechanisms.

Even in a future anticipated hybrid world of humans and intelligent artifacts the relationships between different ethical frameworks and levels of description is necessary. In words of Hongladarom (2008): "*The individual cannot extricate herself from her own specific and fine-grained details of her social and physical environment*."

A similar conclusion comes from Grodzinsky et al. (2008) who also seek to connect LoA of PI with those more everyday ethical issues one is used to: "*at levels of abstraction that are more concrete (i.e., where implementation details are visible)*". This recurring wish for providing more specific examples of connections between IE and classical ethical approaches is the evidence of interest in applying IE analysis.

Focusing on a fundamental level of organization and radically rethinking our relationships with each other and with the world, IE essentially contributes to our ability to understand underlying mechanisms of ethical behavior in networks of humans and intelligent artifacts. The observed progress towards increased distribution of cognitive functions in such systems (Magnani 2007) necessitates application of PI.

1.2.3 On the Artificial Agency and Morality

This article concerns systems of humans and intelligent adaptive artifacts and in the first place the problem of (moral) responsibility distribution. It argues that for all practical purposes, moral responsibility in autonomous intelligent systems is best handled as a regulatory mechanism, with the aim to assure their desirable behavior. "Responsibility" is thus ascribed an intelligent artifact in much the same way as "intelligence" and it is considered to be a matter of degree. We will expect a (morally) responsible artifactual intelligent agent to behave in a way that is traditionally thought to require human (moral) responsibility.

In order to make the point about artificial moral agency (Grodzinsky et al. 2008) adopt concept of Levels of Abstraction and discuss the difference between artificial agents whose behavior is completely defined by their designers, and agents able to learn and adapt, changing their own programs autonomously. They conclude that designers and other concerned stakeholders must maintain responsibility for those artifacts, no matter how autonomous they may be. Actually this conclusion shall not come as a surprise. The question Grodzinsky, Miller and Wolf ask: "*Can an artificial agent that changes its own programming become so autonomous that the original designer is no longer responsible for the behavior of the artificial agent?*" in the perspective of distributed responsibility discussed in detail later on, gets an obvious answer. Such an artificial agent with an artifactual equivalent of "free will" can not be more autonomous than a human within a techno-social system. Even though humans have free will and autonomy, there is a distribution of responsibility in a system.[2]

> *Again: the idea of building moral responsibility into artificial agents is not meant as leaving those agents outside of the techno-sociological control.*

One of the central concepts in this context is the concept of agent. Unlike (Himma 2009) who concludes his essay by the claim that artificial moral agency is possible if it is possible for ICTs to be conscious, in the field of Agent Based Modeling (http://www.scholarpedia.org/article/Agent_based_modeling) agents are supposed to include even much simpler entities. Agent-Based Modeling (ABM) is an individual-based modeling of a phenomenon as a system of interacting agents (actors) such that agents have internal states.[3] Humans may in this context be seen as highly complex agents.

Agents in general may be as simple as cellular automata but may also have random-access memory, i.e. they can interact with the environment beyond concurrent state communication by using memory to save representations of the environment. Members of an agent society can share information and knowledge. Such agents are dynamically incoherent as their next state is not only dependent on the previous state but also on their memory (which keeps the same value until it is accessed). Agent interactions can be local, global or intermediate (small-world network). The system evolves over time, and since agents behave individually in parallel, interactions are generally asynchronous.[4] ABMs are powerful modeling tools which relate Artificial Life, Game Theory and Artificial Intelligence and in this context are useful in the studying of ethics in IE applications.

---

[2] As long as artifacts are under human control, such as GPS devices, we have no problem to follow their command. But what kind of assurance do we need when artifacts with superior cognitive capacities have their own agenda? I believe that we will get successively better insights into that issue as we enhance our own cognitive capacities through distributed cognition in networks of biological and artificial agents.

[3] Internal states are represented by discrete or continuous variables.

[4] In ABM, both time and space can be discrete or continuous.

1.2.4 IE's Constructive/Generative Nature

Enabling computational modeling in IE resembles adding a microscope to medical diagnostic tools. It will not replace a doctor's usual examination of a patient, but will provide a useful complement. The result of the investigation of a patient's health naturally depends on the diagnostic method. So on one level of analysis, the problem might be identified as a high level of leukocytes in the blood. On a higher granularity level, the same problem may appear as an infectious disease. On an even higher, social level, the problem may be characterized as an epidemic and a health-care problem.

Instead of being worried by the fact that different levels of abstraction show different views of the world and give different answers to questions, such as: what is wrong? (the leukocyte number is too high, the patient has an infectious disease, there is a threat of pandemic, etc.), we should be happy with the fact that we finally make explicit a variable always present in every analysis, a variable which otherwise is hidden and often the source of misunderstanding in ethical debate when two parts in a dialogue discuss the problem on different levels of abstraction, without even recognizing that.

In other words, information-centric IE is a complementary and not an alternative approach to traditional ethics. As already pointed out by Floridi (2008a, b, c, d), there is a plurality of possible approaches which may "*complement and reinforce each other, like stones in an arch*."

The strongest side of IE is its focus on the understanding of mechanisms of ethical behavior on a conceptually more fundamental level than what conventional ethical approaches usually provide. Instead of assuming that an agent is perfectly well informed and perfectly rational, the modeling of ethical agent systems on the informational level permits studying effects of information communication and processing in networks of agents. It includes effects of imperfect information transmission and how the global behavior of a system changes when agents get distorted information or no information at all, or when an agent itself is not a perfectly rational human but maybe a less cognitively equipped machine/program. The grounds for normativity in such an info-computational system can be studied by simulation models as well.

Rather than providing an automaton for ethical norm generation from available information, IE presents a valuable tool for studying the effects of plurality of ethical choices and network configurations reflecting the macro-ethical character of IE.[5] As in general with different levels of abstraction, the answers from micro- and macro-perspectives are not necessarily identical.

Being especially suitable for the analysis of artificial multi-agent systems (MAS), Information Ethics may be expected to be at least a useful framework in supporting

[5] Unlike the micro-ethical level where one considers what an individual should do, at the macro-ethical level the question is what macro-systems, such as political institutions, corporations or professional organizations, should do.