

“Understanding Information Ethics: Replies to Comments”

Luciano Floridi^{1,2}

¹Research Chair in Philosophy of Information and GPI, University of Hertfordshire; ²Faculty of Philosophy and IEG, University of Oxford.

Address for correspondence: Department of Philosophy, University of Hertfordshire, de Havilland Campus, Hatfield, Hertfordshire AL10 9AB, UK; l.floridi@herts.ac.uk

Foreword

Information Ethics (IE) is a lively area of philosophical research that has been attracting an increasing amount of interesting work. A recent special issue of *Ethics and Information Technology*, edited by Charles Ess (“Luciano Floridi’s Philosophy of Information and Information Ethics: Critical Reflections and the State of the Art”, *Ethics and Information Technology* 2008 (10.2-3)), and two issues of the *APA Newsletter on Computers and Philosophy*, edited by Peter Boltuc (Spring 2008, 7.2 and Fall 2008, 8.1) testify to the development of several lines of critical investigations of some of the initial ideas and arguments that I have tried to articulate and support during the past decade. Indeed, I take it as a good omen that this article, containing the replies to the articles published in the *APA Newsletter*, appears exactly ten years after “Information Ethics: On the Theoretical Foundations of Computer Ethics” (*Ethics and Information Technology* 1999 (1.1), 37-56), a work that I consider the starting point of my research in the field.

Ten years of sustained, international research of high standards is a long time in any academic field. In Information Ethics this is even more so, given the fast-paced and radical transformations involving information and communication technologies (ICT). As a result, IE has certainly widened its scope. It now interacts with many other ethical fields, from business ethics to environmental ethics, from medical ethics to the ethics of nanotechnologies, from the ethics of cyberwar to the ethics of e-research. But IE has also deepened its conceptual insights. These involve now dialogues with other philosophical and ethical traditions, such as Platonism, Neo-Platonism, Stoicism, Spinozism, Deontologism, Consequentialism, Contractualism and Buddhism, as well as analyses and discussions of metaphysical, epistemological and logical topics, from digital physics to the method of levels of abstraction, from structural realism to the philosophy of information. In this rich and varied context, informed and reasonable debates and disagreements are welcome. Not only are they clear

evidence of a healthy market of ideas, open to different and sometimes contrasting views; they above all can foster our understanding and help to guide our sound judgements. The interested reader will find plenty of such discussions in the comments and in my replies to them.

A final explanation, before leaving the floor to the debate. In writing the following replies, I have tried to see points of convergence but I have also tried to avoid repetitions, so I am afraid the reader will get a full picture only by browsing through the whole collection of replies and some of the texts to which I have referred there. In order to facilitate the reader's task, I have referred only to texts of mine that are freely available online from my website: www.philosophyofinformation.net.

By way of conclusion, let me say that I am very grateful to Peter Boltuc, and to all the colleagues who took their time to write their articles, for this opportunity to acknowledge that much more and better work needs to be done in IE before one may justifiably claim to have a full outline of the theory. My hope is that I have been able to show that such work is worth all our intellectual efforts.

Reply to Vaccaro's Information Ethics as Macroethics: Perspectives for Further Research

Vaccaro's article identifies, with insight, several questions that IE should rightly be called upon answering. They may be reasonably summarised under three headings:

- a) **applicability**: whether IE can help in how we deal with everyday, moral issues.
- b) **inclusivity**: whether IE's treatment of patients and agents as informational entities can improve our analysis and understanding of the full spectrum of applied ethical problems.
- c) **extensibility**: whether IE's ethical principles may be applied to material (as opposed to digital) contexts.

These are very valuable suggestions for further research. They point towards issues that are challenging and complex. So it would be preposterous even to try to outline here how IE may deal with them fruitfully. What they require is actual work in the field, not further theorising or a quick and dirty fix. Nevertheless, they may benefit from a basic clarification, which is brief enough to be provided in the short space available for this reply. It is to be hoped that it may further the debate.

IE is sometimes understood as if it defended the view that artificial, digital or informational realities should also be taken into account by the ethical discourse, besides human and biological agents and patients. This is a reasonable but slightly mistaken view. The thesis actually advocated by IE is significantly different, and perhaps even more radical (so, to a sceptic, it will look even less credible). It consists in arguing for a change in the level of abstraction at which the ethical discourse may also be fruitfully developed. IE fosters a development from biocentrism to ontocentrism, where the latter is expressed in terms of an informational metaphysics. To put it simply, according to IE, the effort to be made consists not merely in adding new agents and patients to the list of already ethically qualifiable entities, but to interpret all agents and patients informationally, thus including humans, animals, social agents and engineered entities as well. It is an extension achieved not in terms of addition, but in terms of modification of the interface through which we analyse moral interactions. It follows that the three fundamental issues of applicability, inclusivity and extensibility should really be answered after the following question: does it help to adopt an information-based, metaphysical approach to ethics? As I anticipated, IE defends a firm answer in the positive, but the latter only points in the direction in which further research should be developed.

Reply to Sullins' So You Say You Want a Re-Ontological Revolution

Sullins' article has many merits, but three of them seem to be outstanding. I shall list them in no particular order.

First, the article finely captures the environmental nature of IE. This is an uncommon feature among other interpreters of IE. Second, it is convincing in pointing out the difference between IE's philosophy of technology and some of its catastrophist, apocalyptic alternatives. Third, it stresses a crucial difficulty. It is on the latter that I wish to comment briefly.

To put it simply, Sullins ask whether *physis* and *techne* may be reconcilable. The question does not have a predetermined answer, waiting to be divined. It is more like a practical problem, whose feasible solution needs to be devised. With an analogy, we are not asking whether two chemicals could mix but rather whether a marriage may be successful. There is plenty of room for a positive answer, provided the right sort of effort is made.

It seems beyond doubt that a successful marriage between *physis* and *techne* is vital and hence worth all our efforts. Our societies increasingly depend upon technology to thrive, but they equally need a healthy, natural environment to flourish. Try to imagine the world not tomorrow or next year, but next century, or next millennium: a divorce between *physis* and *techne* would be utterly disastrous both for our welfare and for the wellbeing of our habitat. This is something that technophiles and green fundamentalists must come to understand. Failing to negotiate a fruitful, symbiotic relationship between technology and nature is not an option. Fortunately, a successful marriage between *physis* and *techne* is achievable. Admittedly, much more progress needs to be made. The physics of information can be highly energy-consuming and hence potentially unfriendly towards the environment. In 2000, data centres consumed 0.6% of the world's electricity. In 2005, the figure had raised to 1%. They are already responsible for more carbon-dioxide emissions per year than Argentina or the Netherlands and, if current trends hold, their emissions will have grown four-fold by 2020, reaching 670m tonnes. By then, it is estimated that ICTs' carbon footprint will be higher than aviation's. (Source: *The Economist* May 22nd, 2008). The good news are that ICTs will also help "to eliminate 7.8 metric gigatons of greenhouse gas emissions annually by 2020 equivalent to 15 percent of global emissions today and five times more than our estimate of the emissions from these technologies in 2020" (Source: McKinsey's *Information Technology Report*, October 2008, "How IT can cut carbon emissions", by Giulio Boccaletti, Markus Löffler, and Jeremy M. Oppenheim). This positive (and improvable) balance leads me to a final comment.

The greenest machine is a machine with 100% energy efficiency. Unfortunately, this is equivalent to a perpetual motion machine and the latter is simply a pipe dream. However, we also know that such impossible limit can be increasingly approximated: energy waste can be dramatically reduced and energy efficiency can be highly increased (the two processes are not necessarily the same; compare recycling vs. doing more with less). Often, both kinds of processes may be fostered only by relying on significant improvements in the management of information (e.g. to build and run hardware and processes better). So here is how IE reinterprets Socrates' ethical intellectualism: we do evil because we do not know better, in the sense that the better the information management is the less moral evil is caused. With a proviso though: some ethical theories, especially in the Christian tradition, seem to assume that the moral game, played by agents in their environments, may be won absolutely, i.e. not in terms of higher scores, but by scoring perhaps very little as long as no moral loss or error occurs, a bit like winning a football game by scoring only one goal as long as none is received. It seems that this absolute view has led different parties to underestimate the importance of successful compromises (imagine an environmentalist unable to accept any technology responsible for some level of carbon-dioxide emission, no matter how it may counterbalanced this otherwise). The more realistic and challenging view from IE is that moral evil is unavoidable, so that the real effort lies in counterbalancing by more moral goodness.

Information and Communication Technologies (henceforth ICTs) can help us in our fight against the destruction, impoverishment, vandalising and waste of both natural and human (including historical and cultural) resources. So they can be a precious ally in what I have called elsewhere synthetic environmentalism or e-nvironmentalism. We should resist any Greek epistemological tendency to treat *techne* as the Cinderella of knowledge; any absolutist inclination to accept no moral balancing between some unavoidable evil and more goodness; or any modern, reactionary, metaphysical temptation to drive a wedge between naturalism and constructionism by privileging the former as the only authentic dimension of human life. Sullins is right in pointing out that the challenge is to reconcile our roles as agents within nature and as stewards of nature. The good news are that it is a challenge we can meet.

Reply to Buchanan's Discursive Explorations in Information Ethics

Buchanan has provided an impressive commentary on some key aspects of IE that I have tried to clarify. I doubt that I could add anything more to her outstanding insights. So, in this brief comment, I shall limit myself to highlight a basic and rather simple thesis, which may help to frame our dialogue.

As I have argued in “Artificial Intelligence’s New Frontier: Artificial Companions and the Fourth Revolution” (*Metaphilosophy*, 2008, 39.4/5, 651-655), the informational turn we are experiencing may be described as a fourth revolution, in the process of dislocation and reassessment of humanity’s fundamental nature and role in the universe. We are not immobile, at the centre of the universe (Copernican revolution), we are not unnaturally separate and diverse from the rest of the animal kingdom (Darwinian revolution), and we are very far from being Cartesian minds entirely transparent to ourselves (Freudian revolution). We do not know whether we may be the only intelligent form of life. But we are now slowly accepting the idea that we might be informational entities and agents among many others, significantly but not dramatically different from smart, engineered artefacts. I called this Turing revolution the fourth revolution. It seems that, in view of this important change in our self-understanding, and of the sort of ICT-mediated interactions that we will increasingly enjoy with other agents, whether biological or artificial, the best way of tackling the new ethical challenges posed by ICTs may be from an environmental approach, one which does not privilege the natural or untouched, but treats as authentic and genuine all forms of existence and behaviour, even those based on artificial, synthetic or engineered artefacts. This is one of the fundamental theses defended by IE.

Reply to Chopra's Comment on "Understanding Information Ethics"

Chopra's article contains a list of questions that are not rhetorical, but reasonable requests for further work. He is correct in indicating both the need of such a large theoretical investment and some of the important areas where it should be made. Of course, Chopra's are also the sort of requests that it is fair to make but equally fair to meet in a context different from the present one. It would be simply impossible to answer here any of the questions that close his article, even very succinctly. However, a much more modest goal seems achievable, and that is to clarify a potential source of confusion. To put it briefly, IE defends an ontological position, not a socio-economical or technological analysis of our current society (the point is well captured by Terry Bynum in his article on my work, entitled "Toward a Metaphysical Foundation for Information Ethics", which appeared in the following issue of the APA Newsletter, Fall 2008, 8.1, see my reply below). The ontological position in itself and the possibility of defending it may both be subject to criticism, but these are different issues. Let me clarify why.

Chopra writes that "The infosphere's contours, then, continue to be shaped by a variety of forces, not all of which break down barriers to information flows". There is a risk here that the infosphere might be interpreted not as a metaphysical concept (i.e. another way of speaking of Being), but as some socio-economic or geo-political space. That would be a mistake. This is not to say that reality (in the metaphysical sense) cannot be shaped by a variety of forces, but that such forces are not those identified by Chopra. One must agree that "It is socio-economic circumstances that permit the ubiquity of this panoply of goods" but also acknowledge that this has nothing to do with the infosphere becoming our ecosystem. Imagine criticising Aristotle's metaphysics of substance in terms of Greek economic conditions in the fifth century BC. The real point at stake is whether we are modifying our metaphysical outlook, from a materialist one, in which physical objects and processes still play a key role, to an informational one, in which objects and processes are dephysicalised, typified and perfectly clonable, the right of usage is perceived to be at least as important as the right to ownership, while the criterion for existence is no longer being immutable (Greek metaphysics) or being potentially subject to perception (modern metaphysics) but being interactable. If all this seems a bit too "philosophical", let me try to provide an illustrative example, following Chopra's approach.

Despite some important exceptions (e.g. vases and metal tools in ancient civilizations or books after Guttenberg), it was the industrial revolution that really marked the passage from a nominalist world of unique objects to a Platonist world of types of objects, all

perfectly reproducible as identical to each other, therefore epistemically indiscernible, and hence pragmatically dispensable because replaceable without any loss. Today, we find it natural that two automobiles may be virtually identical and that we are invited to buy a model rather than a specific “incarnation” of it. Indeed, we are fast moving towards a commodification of objects that considers repairment as synonymous of replacement, even when it comes to entire buildings. This has led, by way of compensation, to a prioritization of *branding*, a process compared by Naomi Klein to the creation of “cultural accessories and personal philosophies” (*No logo*, Canada: Random House, 2000) and of *re-appropriation*: the person who puts a sticker on the window of her car, which is otherwise perfectly identical to thousands of others, is fighting an anti-Platonic battle. The information revolution has further exacerbated this process. Once our window-shopping becomes Windows-shopping and no longer means walking down the street but browsing through the web, the problem caused by the dephysicalization and typification of individuals as unique and irreplaceable entities starts eroding our sense of personal identity as well. We become mass-produced, anonymous entities among other anonymous entities, exposed to billions of other similar informational organisms (inforgs) online. So we self-brand and re-appropriate ourselves in cyberspace by blogs and facebook entries, homepages, youtube videos, and flickr albums. We use and expose more information about ourselves to become less informationally indiscernible. We wish to maintain a high level of informational privacy almost as if that were the only way of saving a precious capital that can then be fully invested by us in order to construct ourselves as discernible individuals. Now, what I’ve been arguing in the past is that processes such as the one I have just sketched are part of a far deeper metaphysical drift caused by the information revolution. I have also argued that IE is the sort of ethics we need to develop if we wish to tackle the moral challenges posed by such profound transformations. I believe Chopra and I agree that it is a worthwhile investment, perhaps even a necessary one, given how the world is quickly changing under our eyes. We need to develop a robust Information Ethics. I agree that much work still needs to be done, but that can only be intellectually exciting.

Reply to Terrell Ward Bynum's Toward a Metaphysical Foundation for Information Ethics

If I were to point out only one, major merit of Bynum's article, I would certainly choose its perceptive interpretation of IE as a metaphysical position. Bynum is absolutely right in framing the debate in terms of a tension between Wiener's more Aristotelian and materialist foundation for information ethics, and my approach, which he correctly describes as more Platonic-Spinozian and, I would add, idealist (see Soraj Hongladarom, "Floridi and Spinoza on Global Information Ethics", *Ethics and Information Technology*, Volume 10, Numbers 2-3, September, 2008). The gap between Wiener and me, however, may be less wide than it seems. For what I have been arguing is that his metaphysics might be too restrictive, not wrong. Even Wiener seems to accept this point sometimes. True, he repeatedly *asserted* that he conceived information exclusively in terms of Shannon's theory, but even the quotations provided by Bynum show that he did not have to limit himself to such a theoretical frame. When Wiener famously described human beings as "patterns that perpetuate themselves" (*The Human Use of Human Beings*, p. 96) the patterns in question may by analogue and continuous, "persistent information patterns" that have little to do with Shannon's concept of information. This would be good news, because Shannon's information entails a view of the ultimate nature of reality as necessarily discrete and possibly deterministic (this is also known as digital physics). This neo-Pythagorean ontology is hardly tenable nowadays, and in a recent article I have argued that it is actually mistaken ("Against Digital Ontology", forthcoming in *Synthese*). Instead, and more constructively, it is possible to show that a concept of information as relational patterns is much more satisfactory, and provides IE with a structural ontology that is more successful philosophically and more easily reconcilable with our current scientific knowledge ("A Defence of Informational Structural Realism", *Synthese*, 2008, 161.2, 219-253). So, in the rest of this reply, I shall illustrate what informational structural realism (ISR) amounts to with the help of some classic analogies.

At least since Plato's images of the line and of the cave, philosophers have often relied on spatial analogies to explain their theories. References to rooms are particularly popular. Sextus Empiricus thought that, in our epistemic pursuits, we are like people in a dark room, searching for gold or shooting at a target: no matter how long the search or the shooting proceeds, it is pointless because in principle there is no way to establish whether any of us has found a nugget or hit the mark (*Outlines of Pyrrhonism* I.52 and II.325). Turing used different rooms for his famous test. And Searle devised a Chinese room for his

counterexample. I shall rely on their examples and suggest a double box analogy to illustrate ISR. But first, a final bit of terminology.

In software engineering, *black-box* refers to a test-design method that focuses on testing functional or behavioral requirements of a program. The methodology treats the *analysandum* as a completely opaque and closed system, avoiding using explicit knowledge of its internal nature or structure to understand the way it works. The opposite methodology is known as *white-box* test design. This allows one to “look inside” the system, and it focuses on using specific and detailed knowledge of the program code to guide the selection of test data. A *grey-box* approach is one that allows only a partial view of the internal properties of the system.

According to ISR, any white-box approach to reality is excluded in principle, given the fact that our knowledge is always mediated by some levels of abstraction (LoA). Forget about getting out of Plato’s cave. There is no God’s eye perspective from without. We look at the world as if we were in Sextus’ dark room. This is the first box. We are inside it, but our goal is not mimetic, nor are our (often *causal*) *interactions* with the furniture in the room unidirectional, as Sextus assumed. Unlike Sextus’, ours is only a grey-box. In the best cases, it allows the reconstruction of the structural properties relating the furniture of the room, i.e., our informational objects. These are our second kind of boxes. As in Turing’s test, they are black-boxes, not directly knowable, but “epistemically interactable” through LoAs. Sometimes, we can indirectly capture their nature by observing their behaviour and mutual interactions, but we do not know their intrinsic (Kant would have said *noumenal*) properties. How we relate them and use them to build other black-boxes is our responsibility. This is the right context in which to talk about a demiurgic power. ISR takes our epistemic goal to be *constructionist* (mind, not *constructivist* in any psychologistic or “sociologistic” sense), not mimetic. Knowledge is not a matter of either (a) discovering and describing, or (b) inventing and constructing, but of (c) designing and modelling reality, its features and behaviours into a meaningful world as we experience it. And one may design and model successfully even in the dark. Intelligibility is the name of the epistemic game, and humanity tries to achieve it at any cost, even when this means distorting reality in the most absurd way, from the conception of a flat earth placed at the centre of the universe to the interpretation of natural forces and events as anthropomorphic divinities or to the assumption of calories, phlogiston and luminiferous ether. Since we wish to devise an intelligible conceptual environment for ourselves, we do so not by trying to picture or photocopy whatever is in the room (mimetic epistemology), but by interacting with it as a *resource* for our semantic tasks, interrogating it

through experience, tests and experiments. Reality in itself is not a source but a resource for knowledge. Structural objects (clusters of data as relational entities) work epistemologically like *constraining affordances*: they allow or invite certain constructs (they are *affordances* for the information system that elaborates them) and resist or impede some others (they are *constraints* for the same system), depending on the interaction with, and the nature of, the information system that processes them. They are exploitable by a theory, at a given LoA, as input of adequate queries to produce semantic information (the model) as output. Transforming constraining affordances into semantic information need not be a metaphysically violent business (as Bacon thought it might), if reality in itself is indeed indeterminate or if we are ready to be led by it insofar as it is determinate. From this perspective, semantic concerns (most importantly reference, representation and truth) belong to the relation among models, that is, among outcomes of LoAs (Kant's phenomenal world of experience), not to the relation between models and reality in itself.

It turns out that we are like Turing's interrogator, since the model of investigation is erotetic: we have indirect (LoA-mediated) access to reality and can query it as a database. Bacon and Galilei shared a similar view. But since our task is not to find out who is who, we resemble Searle in his Chinese room: we get the data input on one side and output semantic information on the other. The difference, in this case, is that we have some understanding of the rules of the game. It makes little sense to ask whether the factual information we gain resembles its source. The Parthenon is as concrete and objective as anyone may wish it to be, but it does not represent marble. Knowing reality is interpreting it constructively, not portraying it passively. Cooking provides a better analogy than photocopying.

So the basic idea behind ISR is quite simple: black-boxes inside a grey-box. The last specification to be added is that these qualifications are LoA-dependent, in the same way as the distinction between being a system and being a component or unit of a system is. A black-box may be opened, but opening it transforms it into a grey-box, in which more black-boxes may be found. Whether *ad infinitum* we simply cannot know. It might be Russian dolls (informational objects) all the way in.

Reply to John Barker's Too Much Information: Questioning Information Ethics

Let me begin with a praise: the article captures perfectly well what other, less acute interpreters often fail to grasp, namely the fact that, when I speak of informational objects, I do not mean news, emails or encyclopaedic entries, but entities understood informationally. With a Quinean slogan: to be is to be an informational entity. With a historical comparison, one may think of Leibniz's monads, or Berkeley's idealistic ontology. It is therefore somewhat puzzling that, in the second half of the article, much space is devoted to a digression, namely the quantification of information Shannon-wise. The reader is referred to Bynum's article and my reply to it for a proper understanding of why the mathematical theory of communication represents the wrong approach. Barker seems to agree. He acknowledges that "overall complexity, or quantity of information, is a poor measure of intrinsic moral worth. Now this conclusion, even if true, may not appear to be terribly damaging to Information Ethics, as the latter embodies no specific theory of how to measure moral worth". This is precisely why we should not be distracted by it. However, the article continues "It may simply be that some other measure is called for". Or perhaps none at all, which is the alternative I would prefer. Shannon's and other similar mathematical approaches to quantitative data, their probability distributions and so forth (see my "Information" in the *Blackwell Guide to the Philosophy of Computing and Information*) are simply irrelevant here. We should not get lost behind false promises of numbers and formulas. When we speak of more or less evil or moral goodness, or higher or lower degrees of flourishing and well-being, these are qualitative assessments that require practical wisdom or *Phronesis*, not a pocket calculator. They may be made more precise by some formalization, but cannot be seriously quantified.

The article continues and concludes with an important challenge, which I shall discuss at the end of this reply. Before, let me stress that the article is correct in stating that, because of my informational metaphysics, I have argued in favour of an expansion of the circle of entities that have some (possibly very minimal and often easily overridable) moral claim on us, in order to include in it every manifestation of Being. Here, however, I would like to rectify a potential source of misunderstanding. According to the article, "Floridi's central claim [...], while fascinating, certainly runs counter to most moral theories that have been proposed". Without starting the usual game of who said what and whose philosophical ancestors are more or less respectable, the idea that Goodness and Being (capitals meant) might be two sides of the same concept, as Evil and Non-being might be, is actually a classic position. Perhaps the objection could rather be that I am defending something hardly new.

Without disturbing Eastern traditions within Buddhism, Hinduism or Shinto – which I understand attribute intrinsic value both to sentient and to non-sentient realities – the reader sufficiently acquainted with the history of Western philosophy will recall that many classic thinkers, such as Plato, Aristotle, Plotin, Augustine, Aquinas and Spinoza, have elaborated and defended in various ways the fundamental equation highlighted above. For Plato, for example, Goodness and Being are intimately connected. Plato's universe is value-ridden at its very roots: value is there from the start, not imposed upon it by a rather late-coming new mammalian species of animals, as if before evolution had the chance of hitting upon *homo sapiens* the universe were a value-neutral reality, devoid of any ethical worth. By and large, IE proposes the same line of reasoning, by updating it in terms of an informational ontology, whereby Being is understood informationally and Non-being in terms of entropy. Note that this is not a defence of IE but an explanation. Although being in the company of Plato or Spinoza, for example, might be reassuring, it is not an insurance against being mistaken. But it is a rectification of the incorrect remark that IE stands rather alone in its defence of what might be called *axiological ecumenism*.

My next comment concerns the statement that “the main rationale Floridi provides [for the axiological ecumenism seen above] seems to be an argument from precedent”. This unfortunate oversight is probably due to a fault of mine. In the article discussed by Barker I provide only a very general outline of IE. So let me take advantage of this opportunity to redress the situation. The reader interested in what I have actually argued may wish to check “On the Intrinsic Value of Information Objects and the Infosphere” (*Ethics and Information Technology* 2003 (4.4), 287-304) and “Global Information Ethics: The Importance of Being Environmentally Earnest” (*International Journal of Technology and Human Interaction*, 2007, 3.3, 1-11). The following should be taken as a mere sketch of the basic line of reasoning.

The actual argument seeks to establish that entities deserve respect because they have intrinsic value, where this causal explanation leads to the interesting problem whether non-sentient entities too may have some (perhaps very) minimal, (perhaps easily) overridable but still intrinsic value. Without rehearsing the whole discussion, I agree that the answer here may be difficult to grasp, as it requires a mental frame rather different from the one any anthropocentric ethics has trained us to adopt. It consists in shifting the burden of proof (a sort of Gestaltic shift) by asking, from a patient-oriented perspective, not “why should I care, in principle?” but “what should be taken care of, in principle?”, that is, whether there is anything that is intrinsically worthless ethically, and hence rightly disrespectful in *this*

particular sense, i.e., insofar as its intrinsic value is concerned (again, something might deserve to be disrespected for other reasons, e.g., instrumentally or symbolically or for the sake of other entities). In short, one line of reasoning in favour of IE's position (there are others, see references above) is that, because we lack arguments against the intrinsic value of Being in all its manifestations, we are led to expand an environmental approach to all non-sentient beings. The injunction is to treat something as intrinsically valuable and hence worthy of moral respect by default, until "proven guilty". The intuitive idea is that a universe without moral evaluators (e.g. humans) would still be morally valuable, and that an ontologically richer universe would be a morally better universe than an ontologically poorer one.

At this point, a standard move is to be told that there is at least one argument against accepting the intrinsic value of Being in all its manifestations: it would be morally too demanding. But the objection is only apparently convincing. Age and experience teaches that there is probably nothing more difficult than living a morally good life. Imagine asking what it takes to win a gold medal at the Olympic Game, and then objecting that the answer cannot possibly be correct because it would require too much effort, or too many advanced skills, or unusual capacities and gifts. The reply might be that this is not a good analogy, because a moral life cannot be as difficult to achieve as a gold medal at the Olympic Game. And it is exactly this reply that unmasks a deeper problem. For it shows that what is at stake is not really a supererogatory issue anymore, but rather the mistaken assumption that lies behind the supererogatory argument: that the moral game is a game sufficiently easy to win, and that it must be so because any human being must in principle be able to win it, and that this is the case because some ultimate salvation is at stake, and a game too difficult or even impossible to win would be unfair. This is a very non-Greek but rather Christian silent axiom. It is not Greek, because Greek culture knows too well the meaning of the tragic: the failure of a good will to do the right thing (see "Information Technologies and the Tragedy of the Good Will", *Ethics and Information Technology*, 2006, 8.4, 253-262). Greek eyes do not fear to see life as intrinsically and irremediably unfair and unjust sometimes. The silent axiom is much more in tune with Christian ethics because the latter presupposes a fair Judge and an Ultimate *Redde Rationem* ("Redde Rationem villicationis tuae", "give me an account of your stewardship", Luke 16:2). IE finds a compromise between these two positions by seeking to interpret the morally good life as a matter of differential score, as I have tried to explain in my reply to Sullins: we cannot avoid doing some evil and this is our tragic predicament, but we can still be good agents if we do more good than evil, and this is our heroic chance. We shall

inevitably fail many times, but we could succeed even more times. The question is then: how could we know how to strike such an active balance? This is the way I understand the serious challenge posed by the end of the article.

The question concerns what sort of metrics one could use in order to determine whether entities (understood as informational objects) have more or less significant (or negligible) moral value and hence how one could live a morally good life. The answer is quite simple: the more an entity contributes to the welfare of the infosphere the higher its status is in the chain of morally respectable beings. Once again, this is a classic position for which I claim little originality. If there is a God, God is the ultimate respectable entity as the source of all entities. A biological virus must, unfortunately be destroyed for the sake of the rest of the environment and its flourishing. Between these two boundaries we can aim to imitate God or run the risk of being worst than a virus. People have managed both.

Nobody could ever argue that a spider's and a human life are equally worthy of respect. Culling, for example, is an ethical duty in environmental ethics. Likewise, the destruction of entities might easily be not only inevitable but welcome in IE. Again, IE is not about respecting a single grain of sand as much as one respects the whole earth full of life or other human beings. It is about placing the threshold below which something is morally disrespectful in itself and rightly so. With a Cartesian analogy, the mistake lies in thinking that, if one argues that all physical things are extended, then one is arguing that they are all of the same size. Of course they are not, and nobody could reasonably argue that they are. To revert to IE, the view that all entities are *at least minimally and overridably* valuable in themselves should not be confused with the view that they all share the same value. As for IE offering little guidance once the moral worthiness of all aspects of Being (axiological ecumenism) is accepted, this would be equivalent to saying that, since environmental ethics is based on the value of life and of the absence of suffering, then it offers little help with real-world issues. The truth is exactly the opposite. Having some general, basic and robust principles in place helps enormously when it comes to dealing with complex, practical matters. We should not fear to respect any form of reality too much, even if this might be a rather difficult task. Rather, as Augustine nicely put it, *dilige, at quod vis fac* (love/respect and do what you wish).

Reply to Edward Howlett Spence's Understanding Luciano Floridi's Metaphysical Theory of Information Ethics

Spence's article contains many valuable insights and suggestions. It also provides some sceptical remarks. In previous replies and many other contexts I have explained why these may be either incorrect or correct but misdirected, as they are not relevant to IE (I refer the interested reader to "Information Ethics: A Reappraisal", an article that contains a foreword to my replies and the replies to a collection of papers edited by Charles Ess and published in "Luciano Floridi's Philosophy of Information and Information Ethics: Critical Reflections and the State of the Art", Special issue of *Ethics and Information Technology*, 10.2-3, 2008). So, instead of indulging in a critique of the criticism, in the short space of this reply I would like to discuss two new points.

The first point is scholarly. In an article that has unfortunately escaped Spence's otherwise thorough analysis ("Global Information Ethics: The Importance of Being Environmentally Earnest", *International Journal of Technology and Human Interaction*, 2007, 3.3, 1-11.), I have explicitly called attention to the Stoic nature of IE. So Spence is partly right when he writes that IE "seems to offer a kind of Stoic Pantheistic Ethics (my [i.e., Spence's] phrase) that endows everything in the Universe with a moral significance and status through a pre-determined divine rational order in which everything is ontologically inter-connected and of which everything forms an ontic part, no matter how big or small." However, I would disagree about the fact that pantheism or religious forms of determinism play any role in IE or in my philosophy in general. I cannot see how anyone could believe that they do. This scholarly point has an important philosophical implication, which I can only summarise here (it is made more fully and repeatedly in several of my writings on IE). IE seeks to break the artificial constraints of what may count as morally valuable. One way in which it tries to escape from such ethical Chauvinism is by showing that there is no good reason to raise any barrier. To put it simply, this is like arguing:

- a) P is the case because it is not the case that $\neg P$.

This is classic, elementary logic. Of course, it is also a way of reasoning that one may not wish to endorse for equally good logical reasons, one only needs to recall Intuitionistic logic or forms of anti-realism à la Dummett. But accepting the logic and not its issuing constrains on the validity of the reasoning is mere inconsistency. Let us now turn to the article. When Spence writes that

- b) “I agree with Floridi that there would be no good reason not to adopt such a higher and more inclusive moral perspective if there were, in fact, good objective and independently grounded reasons for adopting such a perspective”

a logical mistake is made and a crucial philosophical insight goes missing.

The mistake is the following: (b) is not a version of (a) but is rather equivalent to

- c) if there were good Rs (Spence’s “good objective and independently grounded reasons”) to adopt P (the thesis concerning the intrinsic moral goodness of Being, or axiological ecumenism) then there would be no good Rs not to adopt P;

but note that while (a) is formally valid, (c) is not (and not just because it is expressed subjunctively): the premise could well be true, e.g. there could well be good Rs to adopt P, while the conclusion still being false, that is, while there could also be perfectly good Rs not to adopt P. Indeed, this is a very common scenario in our moral lives, which are full of dilemmas: there are often good reasons both to adopt a position, a decision, a course of action, make a choice etc. and balancing, equally good reasons not to do so. This is why I have never supported either (b) or (c), and hence I must turn down Spence’s offer to agree with me on a mistake I have never made.

The missing philosophical insight is connected to (a). One can immediately see that the argument in question is negative or indirect. It consists in reminding historically and showing logically that we have nothing to fear from a holistic attitude towards the value of Being in all its aspects; that it is fine to start from the presupposition that no entity deserves moral disrespect in itself; that anything less than a holistic attitude towards the value of Being would be *prima facie* unjustified. Throughout the article, this crucial point is entirely ignored. That is why Spence keeps repeating that he can find no reasons in favour of P: the reasons are against not P. Since this step is never explicitly grasped nor implicitly followed, unsurprisingly the result is that Spence and I move in very different directions. This is a pity, since Spence seems to be interested in reaching conclusions that are very similar to IE’s.

The second point is more metaphysical: suppose Spence had identified and accepted the previous logical step. What could be the ontological foundation of such an axiological ecumenism? The answer requires the introduction of the concept of *ontic trust*.

Various forms of contractualism (in ethics) and contractarianism (in political philosophy) argue that moral obligation, the duty of political obedience, or the justice of social institutions, have their roots in, and gain their support from a so-called “social contract”. This may be a merely hypothetical agreement between the parties constituting a society (e.g. the people and the sovereign, the members of a community, or the individual and

the state). The parties accept to agree to the terms of the contract and thus obtain some rights in exchange for some freedoms that, allegedly, they would enjoy in a hypothetical state of nature. The rights and responsibilities of the parties subscribing to the agreement are the terms of the social contract, whereas the society, state, group etc. is the entity created for the purpose of enforcing the agreement. Both rights and freedoms are not fixed and may vary, depending on the interpretation of the social contract.

Interpretations of the theory of the social contract tend to be highly (and often unknowingly) anthropocentric (the focus is only on human rational agents) and stress the coercive nature of the agreement. These two aspects are not characteristic of the concept of *ontic trust*, but the basic idea of a fundamental agreement between parties as a foundation of moral interactions is sensible. In the case of the ontic trust, it is transformed into a primeval, entirely hypothetical *pact*, logically predating the social contract, which all agents cannot but sign when they come into existence, and that is constantly renewed in successive generations.¹ The sort of pact in question can be understood more precisely in terms of an actual trust.

Generally speaking, a trust in the English legal system is an entity in which someone (the trustee) holds and manages the former assets of a person (the trustor, or donor) for the benefit of certain persons or entities (the beneficiaries). Strictly speaking, nobody owns the assets, since the trustor has donated them, the trustee has only legal ownership and the beneficiary has only equitable ownership. Now, the logical form of this sort of agreement can be used to model the ontic trust, in the following way:

- the assets or “corpus” is represented by the reality, including all existing agents and patients (the infosphere);
- the donors are all past and current *generations* of agents;
- the trustees are all current *individual* agents;
- the beneficiaries are all current and future *individual* agents and patients.

¹ There are important and profound ways of understanding this *Ur-pact* religiously, especially but not only in the Judeo-Christian tradition, where the parties involved are God and Israel or humanity, and their old or new *covenant* (διαθήκη) makes it easier to include environmental concerns and values otherwise overlooked from the strongly anthropocentric perspective *prima facie* endorsed by contemporary contractualism. However, it is not my intention to endorse or even draw on such sources. I am mentioning the point here in order to shed some light both on the origins of contractualism and on a possible way of understanding the onto-centric approach advocated by IE.

By coming into being, an agent is made possible thanks to the existence of other entities. It is therefore bound to all that already is (the infosphere) both *unwillingly* and *inescapably*. It *should be* so also *caringly*. *Unwillingly*, because no agent wills itself into existence, though every agent can, in theory, will itself out of it. *Inescapably*, because the ontic bond may be broken by an agent only at the cost of ceasing to exist as an agent. Moral life does not begin with an act of freedom but it may end with one. *Caringly*, because participation in reality by any entity, including an agent – that is, the fact that any entity is an expression of what exists – provides a right to existence and an invitation to respect and take care of other entities. The pact then involves no coercion, but a mutual relation of appreciation, gratitude and care, which is fostered by the recognition of the dependence of all entities on each other. A simple example may help to clarify further the meaning of the ontic trust.

Existence begins with a gift, even if possibly an unwanted one. A foetus will be initially only a beneficiary of the world. Once she is born and has become a full moral agent, she will be, as an individual, both a beneficiary and a trustee of the world. She will be in charge of taking care of the world, and, insofar as she is a member of the generation of living agents, she will also be a donor of the world. Once dead, she will leave the world to other agents after her and thus become a member of the generation of donors. In short, the life of an agent becomes a journey from being only a beneficiary to being only a donor, passing through the stage of being a responsible trustee of the world. We begin our career of moral agents as strangers to the world; we should end it as friends of the world.

The obligations and responsibilities imposed by the ontic trust will vary depending on circumstances but, fundamentally, the expectation is that actions will be taken or avoided in view of the welfare of the whole world.

The ontic trust is what is postulated by the approach supported by IE. According to IE, any form of reality (any instance of information/*being*), simply by the fact of being what it is, enjoys a minimal, initial, overridable, equal right to exist (be left alone) and develop (not to be interfered) in a way which befits its nature. Nothing is too humble to deserve no respect at all, i.e. to be rightly disrespectable in itself, not even a rock on the moon. In this way, IE brings to ultimate completion the process of enlargement of the concept of what may count as a centre of a (no matter how minimal and overridable) moral claim, which now includes every instance of Being.

The acceptance of the ontic trust requires a disinterested judgement of the moral situation from an objective perspective (or Level of Abstraction), i.e. a perspective which is as non-anthropocentric as possible. Moral behaviour is less likely without this epistemic

virtue. To argue that any ethical judgement is inevitably human and subjective because it is the result of some human subject's activity is to fall into the fallacy of believing that since a message comes from a given source it must inherit all that source's properties. This is nonsense. An iPod does not play Mac music. Likewise, just because ethics is humanity's business it does not have to be the business of humanity only. Objectivity can be the outcome of a subjective process. The world of mathematics and logic is there as a proof.

By way of conclusion, let me briefly comment on Spence's Information Ethics without Metaphysics. It seems to me to run against several difficulties, two of which are worth mentioning here.

First, purposive or goal-orientated behaviour cannot confer any value to entities which have no purpose or goal. This includes virtually the totality of the universe. At this point, we either have to acknowledge that all the teleological analysis is just another way of speaking about us as the only entities worth moral respect. Or we have to expand the teleological analysis in various ways to make sure that it becomes a blanket that does cover all that we wish it to cover in the first place. In the former case, the position is unsatisfactory as just another case of the old-fashioned anthropocentrism that IE has been trying to improve. In the latter case, the adjustments are equally unsatisfactory not just because they are *ad hoc* but because they end up being increasingly expensive metaphysically: we start having to accept counterfactual cases ("yes, John is just a corpse now but if he were alive he would be able, etc." you get the picture) and end up endowing with purposive or goal-oriented behaviours any manifestation of Being (but what is the purpose of the moon?). But a fully purposeful universe can only be the result of a fully purposing Architect. I remain therefore sceptical about the value of Spence's proposal, which seems to me to be either unfeasible or more demanding, metaphysically, than IE. It is often the case that one philosopher's use of Ockham's razor is another philosopher's chainsaw massacre.