Luciano Floridi

Replies to Broy, Gabriel, Grunwald, Hagengruber, Kriebitz, Lütge, Max, Misselhorn, and Rehbein

Preface to my replies

I am most grateful to Thomas Buchheim, Jörg Noller, the editorial team, and the publisher of the *Philosophisches Jahrbuch* for the remarkable honour of being invited to contribute to the »Jahrbuch-Kontroverse« series. The scholarly attention paid by colleagues to one's own work is the greatest gift one may receive in academia, even more so these days, when we seem to have increasingly less time to study, think, and dialogue. I have replied to the comments in the same alphabetic order in which they appear in the publication. Here, I only wish to add three remarks concerning all of them.

The first is about a regret. I failed to inform readers of the article that it is only an abridged version (ca. 17,000 words) of a book (ca. 65,000 words), already published in Italian (Floridi 2020b) and forthcoming in English in 2022, entitled: *The Green and the Blue – Naïve ideas to improve politics*. As it becomes clearer from the comments and my replies, I believe that many of the justified requests for clarifications, further justifications, terminological definitions and so forth would have been formulated differently if I had warned the colleagues about the nature of the article. The book is written in the same »naïve style« but does one crucial thing that the article is missing: it presents 20 chapters that provide the framework within which the last chapter, the one containing 100 theses (the article provides »only« 69), may be understood more easily. I apologise.

The second remark concerns my gratitude towards the colleagues who took the time to comment on the article. In many cases, the questions, criticisms, and indeed even the misunderstandings contained in the comments will be precious to improve the text of the English

version, which will be in any case further expanded with respect to the Italian edition (I need to add at least one more chapter on digital sovereignty, already outlined here (Floridi 2020a)).

The last remark concerns a commitment: I know that I need to study much better Arendt, Buber, Jonas, and Levinas. I promise to do my homework before the English edition is published. [379]

Reply to Manfred Broy

If you could look at my copy of Broy's comments, you would find way too many passages highlighted. Not just because we agree on many fundamental issues – we do – but because, on many of these issues, Broy shares the rights questions and the insightful comments required to move further and develop our understanding more deeply. The highlighted passages are places where he is asking for more because more is actually needed. In this reply, I shall limit myself to commenting on only a few such passages, but I recommend reading his text carefully and doing the homework he is rightly suggesting.

Broy is correct that some of the conceptual changes we experience today – in particular, think of a shift from a substantialist to a relational ontology – predate the digital revolution:

> Obviously, these changes have started more than 100 years ago, 20 years before Zuse built the first programmable computer. At this time, there was nothing what could be called the digital or digital natives which are today much more related to networks, relations and to sets, and which influence and form the structure of our society. There seems to be a feedback process going on here between changes in the society and those caused by the digital – all run by the »Blue« and in no way by the »Green«. (85)

He is also right in stressing that there is a feedback mechanism in place. I would only add that there is also a mechanism of »realisation« (Floridi 2018): the digital revolution has catalysed, highlighted, and brought into a shared narrative conceptual changes that have long historical roots. Think of the Copernican revolution and its impact on how we conceptualised ourselves, no longer at the centre of the universe. Of course, we were never at the centre of the universe; we just did not know it. The Copernican revolution was the turning point, as it made us scientifically aware of such a peripheral position,

but it was not unprecedented, nor did it get immediate acceptance. The re-conceptualisations brought about by the digital revolution are comparably deep – this is why I have spoken of a fourth revolution, with Turing coming after Copernicus, Darwin and Freud (Floridi 2014) – but they have long historical roots. If anything, the digital revolution has made us vividly and widely aware of such changes, bringing them to cultural maturity and visibility.

Broy is also correct to call attention to China. I could not agree more. There was no space to do so in the article, but I have sought to analyse China's policies elsewhere (Roberts et al. 2021). Like many people, I am concerned by the increasing economic and military power acquired by this autocracy, but I am even more worried by the cultural influence associated with such a power. As Europeans, it would be terrible to jump out of the frying pan of Americanism into the fire of »China-ism«. This leads me to the last comment I wish to share here. Broy rightly highlights that there is not, and indeed there should not be, only one (totalitarian, I would add) human project. Our ethical perspectives can differ significantly and should not be eradicated in the name of »one planet – one people – one human project«. We still have a living memory of the tragic horrors that this way of thinking caused in Europe last century. This is why I argued elsewhere that we should reconsider the [380] modern foundation of our liberal societies in terms of tolerance first and then justice, rather than justice-only (Floridi 2015, 2016d). However, we should also recall that there is much about which we all agree, think for example of the UN Sustainable Development Goals; that there is much on which we are not ready to compromise, and rightly so, think of the protection of human dignity and fundamental human rights, not negotiable, no matter what culture, place, or time we are considering. Pluralism is not relativism; this becomes clear if one looks at design practices. If you search on Google for images of »chair«, you will see that human ingenuity has created a vast number of artefacts, all counting as chairs. Yet all these artefacts have fundamental elements that they share: they are pieces of furniture meant to be for only one person to sit on (if more than one person, then they start looking like a sofa), they have a back (if they do not they are stools), they have some kind of legs (usually but not necessarily four), they typically do not have side support for a person's arms (when they do they are called armchairs), they are not meant to support your legs (those are chaise longue). We

recognise them everywhere, even if they come in a wide variety of styles, materials, functions, sizes, and so forth. Broy remarks that:

> Luciano Floridi is right when he complains about the absence of a human project in our information society – also in Europe. In fact, we do not have a human project for the digital age. (90)

The human project I am talking about is like a chair: there can be an infinite yet bounded number of ways (think of real numbers between 1 and 2) of designing and implementing it, some of which are more successful than others, but just because we are aware of this pluralism, we should not fall into the trap of thinking that it has no boundaries and anything goes: a lamp or a bed are not chairs as a matter of fact, and some poorly designed chairs (for example, too fragile, too expensive, too ugly, too uncomfortable) are chairs nobody wishes to have or use. The same holds true of the variety of human projects we encounter in the history of humanity. So, if I can abuse the analogy, the question I have sought to address in the article and the much longer book is: what is the right chair we need to design and build today? The answer I have defended is: the Green and the Blue. If you think of it, it really is quite obvious.

Reply to Markus Gabriel

I found Gabriel's comment very useful. It shows me where I failed to communicate, and hence being convincing. Therefore, in this reply, I shall try to ameliorate the situation.

Section one provides some justified scolding. I do not refer to some authors, I miss some references, I should have built more links with existing lines of thinking. To my justification, let me say that in this article (and in the book it comes from) I am not interested in the history of ideas. One may argue that I should, and I accept that. But whether it is a feature (for me) or a bug (for Gabriel), the absence of any refer[381]ence to Bruno Latour, for example, should be seen as meant. It is not an oversight. I know, of course, about actor-network theory. I read more essays by my students about it every year than I wish to remember. I meant to avoid it. So, following the comment about the

> astonishing absence of references to already existing relational contemporary social and political ontologies (96)

the question is why I decided not to use such references. There is no space here – nor do I believe it is interesting – to expand on this, but let me just point to a previous comment:

> Despite his recourse to the very idea of a »human project for the digital age«, Floridi seems to be ensnared by a certain post-modern and posthumanist siren song that is a constitutive part of the problem Floridi wants to overcome. (95)

I spent half of my academic life as an analytic philosopher, doing mostly logic, epistemology, and studying the sort of philosophers and philosophies that consider »post-modern« and »post-humanist«, with or without the hyphen (you never know whether the hyphen is meaningful), insulting epithets. I am not proud of it, I repented, I moved to a department of social science to abandon my old faith and try to open my mind, I no longer consider myself an analytic philosopher, but I hope I may be forgiven when conceptual confusion still triggers in me a natural reaction. It is precisely because I stay away from such »post-modernist« and »post-humanist« ways of thinking that I do not link my line of reasoning to them, Latour included. It is a matter of simple coherence (I shall say something more about *post-anything* labels and why I do not use them in my reply to Hagengruber). As for the rest of section one, since I agree with the objections, I must clarify that they are directed to someone else's position, not mine. All my references to mathematics and physics are meant to be mere illustration, not methodological applications or import. In other words, I agree with the following passage:

> I have a series of objections against the idea of grounding a transformation in (social and political) ontology on an analogy with mathematics and natural science, for the objects of (social and political) science cannot be meaningfully modelled in terms of natural science. There is no social vector space and category theory is not capable of getting the kind of qualitative experience into view that is constitutive of »the participant stand-point«, to invoke Strawson's felicitous formulation. (96 f.)

So, this is where I failed to communicate. To my defence, I can only add that I thought it was too obvious to state it. Who in his right mind could believe that »objects of (social and political) science could be meaningfully modelled in terms of natural science« I do not know, certainly not me. I have never been convinced even by Leibniz's *calculemus* or any Carnapian approach, not even when I was

an analytic philosopher, let alone now. So, I am glad to agree with the objections because they are almost all correct but irrelevant. »Almost all« because there is, however, one that is relevant, but luckily, it is incorrect, because based on lack of knowledge of the methodology it discusses: [382]

> If there is a right and a wrong level of abstraction, in what does the rightness consist? It cannot be reduced to »a way of describing the world«, as there are indefinitely many such ways of describing the world. There has to be some set of criteria that help us to decide which of the available modes of description better capture how things really are. (97)

It is well known that the method of Levels of Abstraction (LoAs, the method can be somewhat technical, but for a simple introduction see [Floridi 2016c]) avoids relativism by adding a crucial element missed by the objection: the purpose for which a specific LoA is adopted. Imagine describing a building. You can describe it in various ways, depending on the chosen observables (an unfortunate misname, they are just conceptual variables, nothing to do with »observation«) and hence the LoA adopted: architectural, economic, historical, psychological, social, etc. The objection seems to imply that any LoA will do and hence that one cannot evaluate or judge which issuing description of the system (i.e., which model) is preferable. This is correct, but only if one misses the point that an LoA models a system (the building, in our example) for a purpose, and it is the purpose that enables the comparison and the evaluation. Consider the following example. If Alice's purpose is to know whether something is the same building in terms of its function, the right LoA may indicate that it used to be a hospital, but it is now a school, so the answer is no, it is not the same building, and furthermore an LoA that models the building in terms of its economic value would be incorrect (it would not address the purpose). But if Alice's purpose is to know whether it is the same building in terms of location, e.g., because two people referred to it to give her some instructions on how to get somewhere, then the answer is obviously yes, the two people were referring to the same building, while the economic LoA would still remain incorrect. So LoAs can be compared, in terms of being more or less correct, depending on the purpose, and these can be more or less fitting depending on the questions one is addressing. The real debate is about what the correct LoA is, *given a purpose*, not whether an LoA is possible or not. The temptation is to ask absolute questions, without asking why (what for,

for what purpose) the question is being asked in the first place. But absolute questions, that is, questions devoid of any indication of why they are asked in the first place and hence what LoA would in principle adequately answer them, should be resisted because they only lead to an absolute mess. Theseus' ship, with some pieces changed, is not the same ship if it is a collector asking, but it is the same ship if it is the taxman asking. The reply, that we need to ask the question at the ontological level, is precisely why I insist, as evident in the article and in all my writings, that I would rather maintain some Kantian, sensible approach and hence an epistemological and not an ontological interpretation of the method of abstraction. In philosophy of science, this leads to an information-based structural realism (Floridi 2008) that is »ontologically committed« (in Quine's sense) only in terms of epistemological choices. Of course, anybody is welcome to wonder what the ultimate answers about the intrinsic ontology of *noumena* may be, but as far as I am concerned, I would rather avoid what I believe to be a nonsensical waste of time.

Let me move to another failure in my communication. I think this question well summarises it: [383]

> I wonder why Floridi does not extend his dialectical operation (political abstention is itself a political act etc.) to his own decisions? (100)

I think I did. I also thought that the point I was making was not very controversial, historically speaking, so clearly something went wrong. Let me try again: in general (history and religious texts and practices provide the evidence), even the best kind of religion tends (of course not always, not everywhere) to support a single, often intolerant view of what that human project is; whereas the best kind of ethics tends (of course not always, not everywhere) to be tolerant (ethics texts and practices provide the evidence). If religion wins the battle for hearts and minds, ethics is often at risk (consider just LGBTQ+ rights). One only needs to check what happens in the US or in Iran. If ethics wins, there may be a better chance that religion may be tolerated. Perhaps I am too simple-minded, but it is the dialectic of tolerance that I had in mind. If I recall correctly, it is the reason why John Locke said that one should be tolerant towards everybody but the Catholics because they are so intolerant that, if they were tolerated, they would take over, and that would be the end of the tolerant people. This leads to a famous problem, called at different times the paradox (Popper) or dilemma (Rawls) of toleration: how far is too far? This is not the place

to discuss it, but the reader interested in its analysis and a possible way of resolving it may wish to check (Floridi 2015). To summarise, I do not see the relation between ethics and religion as an opposition because the former can support the latter. In a completely different context, when I was invited by the *Osservatore Romano* (the daily newspaper of Vatican City State) to comment on »Fratelli tutti«, the third encyclical of Pope Francis, I tried to explain this by stressing that – concerning the three theological virtues: faith, hope, and charity (love) – believers, agnostics and atheists can still agree on charity, which can and should unite all of them, even if they hold irreconcilable views about the first two. So, no opposition, but tolerant inclusion of religion by ethics, is what I meant to express.

I better stop here. There are too many other misunderstandings to list them, like the view that I support some EU-centrism (I thought it was clear that I was speaking of an expectation of leadership by example: as in other contexts, the EU should prove its commitments to fundamental values by showing it in practice), or like the suggestion that I may be arguing for the inclusion or expulsion of EU member states depending on political orientations (of course this would be insane, I thought it was clear that I was speaking of some necessary flexibility linked to the most severe violations of human rights, toleration is not without limits, see discussion above, the EU should not be a club one can never be asked to leave no matter what atrocities one commits). I still hope that a more careful and charitable reader will avoid these misunderstandings by reading what I have written. However, I took full responsibility for these shortcomings. The fact that Gabriel's comment is littered with so many misunderstandings means that I inadvertently failed to convey my ideas in a sufficiently clear way. Reassuringly, this means that I can easily agree with the »objections« moved by Gabriel because they fail to address what I meant. And even more constructively, I agree with the last paragraph of his comment, which I finally recognise as a correct summary of some of the points I tried to [384] articulate in the article. Shame on me for such an apparent lack of clarity. As a former analytic philosopher, I promise to try to do better in the forthcoming book.

Reply to Armin Grunwald

The comment by Grunwald reminds me of one of those amazing technological artefacts one can sometimes observe in a museum, where the archaeologist is able to reconstruct and show a whole mechanism and its inner workings from just a few original bits of rusted metal. As I mentioned in the Preface, the article that I published is only a concise, simplified and heavily pruned synthesis of precisely the artefact that Grunwald has managed to reconstruct with remarkable insightfulness and patience. As he writes:

> Luciano Floridi's »grand narrative« rather needs a monographic book project as a suitable form. (117)

Indeed, and the book is available, only in Italian but I hope soon in English (and in a more expanded version). As I anticipated in the Preface, I regret having failed to inform the readers of the article that the latter is not the whole Netflix series, so to speak, but more like a trailer. Once this is clear, it is impressive how well Grunwald has guessed the rest of the narrative from the available fragments. It follows that many of the valid requests made by Grunwald are entirely justified, about terminological clarifications, links with parts that seem otherwise only connectable yet not connected, supporting arguments, and so forth. By way of clarifying what this may mean, let me offer one specific example by relying on the chapter in which I explain what I mean by »mature information society«.

We are so familiar with talk of »*the* information society« that we sometimes forget that there is no such thing, but rather a multitude of societies, different from each other, some of which may qualify as information ones in different ways and degrees. So, we should really speak of »information societies« without a »the« but with an »s«, and ensure that our generalisations are not so generic as to apply to all of them, while obliterating any salient distinction. Just to be clear, there is always a level of abstraction at which something is like anything else: the moon is like your umbrella, which is like a pizza, because they are all individual objects that exist and look round, for example. The point is not being smug about one's own acrobatic equations (x is like y, which is like z) but being critical in checking whether the level of abstraction at which the equation is drawn is fruitful to fulfil the purpose that one is pursuing. All this should clarify why, once we have many information societies that are all different from one another, it

still makes sense to compare them in terms of relevant criteria and why, more specifically, it is essential to understand what it means for an information society to be more or less mature than others. In the article (and in the book), I refer to »maturity« as a matter of people's expectations, not technological or economic development, let alone ethics or civilisations. Let me explain this using an analogy.

When you are in a hotel in Paris, you rightly expect the water in the bathroom to [385] be drinkable because France is a »water mature society«. In fact, you do not even think about it. There is no need for the hotel to advertise the safety of its water, nor for you to ask at the reception whether the water is drinkable. France is a »water mature society« not just because of its water system, but because people living there treat drinkable water as something ordinary, non-informative, a matter of fact that lies in the background. It is part of life, of what anyone implicitly and unreflectively expects the water to be like in Paris. At the same time, we all know that drinkable water is not a trivial matter. There are hundreds of millions of people who do not have access to safe water. So, if you take a more adventurous holiday in an unfamiliar place, your expectations change. It becomes normal to inquire whether it is safe to brush your teeth with the water from the tap. Clearly, expectations change contextually. They are a good way to gauge the maturity of the society in which one lives. The formula is simple: if the occurrence of a feature F in a society S is no longer informative, but it is rather its absence that it is, then S is F mature. According to this interpretation, we are already living in mature information societies in some corners of the world. In such corners, we expect, as a matter of course, to be able to order any kind of goods online, to pay for them digitally, to be able to exchange any sort of contents on the web, to search for any question and find any bit of information, to use services, stream entertainment, and so forth, and all this 24/7, seamlessly, quickly, and reliably, without asking anymore whether it is possible, or being astonished that it is. We realise we live in a mature information society only when such expectations are unfulfilled. Once we analyse information societies in terms of their members' unreflective and implicit expectations, comparable to having drinkable water in Paris, then we can switch from quantitative to qualitative assessments, and consider some significant consequences. This is what I do more extensively in the book and partially in the article (the interested reader may wish to

check (Floridi 2016b), which is the original text in English of which the chapter in the Italian book is a revised version).

The exercise could be repeated, but I hope the previous example shows how brilliant the reconstruction and critical assessment provided by Grunwald is.

Reply to Ruth Edith Hagengruber

I have learnt much from Hagengruber's comment. This quotation well summarises one of the lessons I enjoyed most:

> In this model [Bruno's], a unit is not seen as an immobile entity, as a part in a system of wholes, but as the capacity to entail differences, the more, the wider, the better and the stronger. The turn from a part-whole driven ontology to a perspective of things as objects of information began. (123)

It is a good reminder, to myself included, that new ideas are often old ideas that did not make it to the surface of our popular culture or academic discourse before. And as someone who has done research on the Renaissance and the transmission of knowledge (Floridi 2002), I agree entirely with Hagengruber. I think that the lesson [386] she outlines has remained largely unapplied, and that we still reason, in everyday life and in socio-political contexts, way too much in Aristotelian terms (small blocks making bigger blocks, the »Lego-like« ontology that I criticise in the article) and not enough in terms of relations, nodes, and connectedness. If this is more a Platonist tradition, I can only be delighted, having always been more of a friend of Plato's than of Aristotle's. There is more to learn from her text, for example the insightful link she highlights between the article and another work of mine in which I discuss the fourth revolution in our conceptual displacement (Floridi 2014) from the centre of the universe, of the animal kingdom, of the mental space and now of the infosphere (Copernicus, Darwin, Freud, Turing). But let me make one contribution to our dialogue lest I may appear too lazy. Hagengruber frequently refers to »post-humanism«, for example here:

> The post-humanism of *the Green and the Blue* is only another step in the series of »lost uniqueness« and domination. (125)

She does so interestingly, but I have avoided using the same terminology in my own work for the same reason I avoid presenting a human

project for the twenty-first century in terms of Enlightenment or a new Renaissance. It is not just because these are periods that we have glorified through historical narratives that have been particularly selective and well-edited (Athens forced Socrates to commit suicide; the Enlightenment is also the guillotine; the Renaissance is also when Bruno was burned at the stake, etc.). This is the case, but perhaps it might still be fine, as long as we all know what we are engaging in an exercise of selective memory. It is also not because all these periods are »white male« periods, just to use Hagengruber's recurrent expression, and anyone unhappy with the qualification may also be reluctant to adopt them. Instead, it is because I am convinced that we need to understand our present and possibly design our future more autonomously, learning the lesson from those periods, which should teach us precisely the opposite of any post-anything approach: more intellectual independence. The Renaissance did not define itself post-medievalism, and so forth. We need to find our own voice, not simply appropriate our ancestors' or, even worse, define ourselves in terms of what we are not, post-this or post-that. I do not have a good suggestion to replace these labels, and this is a shortcoming of the point I am making here, for one may argue that, in the absence of any better conceptualisation, we may as well fall back to a good one. But, at least in my case, I would like to leave the conceptual space empty, and feel the pain of the absence, rather than filling it with post-humanism, or neo-Enlightenment or post-modernity or… any other ready-made label that invites us to be conceptually lazy and enjoy repetition, rather than risk novelty. So, I am happy to follow Hagengruber in her analysis and appreciate the terminological nuances but, when it comes to my own conceptual design, I am ready to feel that unpleasant feeling of a missing concept lingering on the tip of my tongue and yet still escaping a complete formulation. *Almost* like some kind of »post-humanism«, … but not really. [387]

Reply to Alexander Kriebitz, Christoph Lütge, and Raphael Max

The comment by Kriebitz, Lütge, and Max is an excellent example of clear and substantive thinking (the valuable distinction between changes of degree and changes of type is a good case in question, where some constructive disagreement could bear fruit). There is much to

praise in both the content and the reasoning, but I shall limit myself here to stress the value of an »order ethics approach«.

I learnt about order ethics from Lütge himself a long time ago, and I have been convinced ever since that it may be a great companion of travelling for the explorations in which I engage. In case, it is only my fault that I have not relied on it more. The following quotation shows why:

> Differing from first order approaches, which take a specific moral framework such as utilitarianism or deontology for granted, second order approaches – of which order ethics is one – are about solving conflicting statements on morality between first order approaches and solving situations in which conflicting normative expectations confront individuals. From the perspective of order ethics, the main purpose of ethics is to define the normative foundations of societies under the condition of moral pluralism and to elaborate principles and structures that overcome failures in cooperation. Different from virtue ethics or deontological approaches, norms derive here from the mutual consent of individuals, with the ultimate goal of reaching mutual improvements by cooperation. (138)

Any reader of the article will see that this is very close to what I have discussed there and in other contexts (Floridi 2016a, 2017) in terms of »infraethics«. By way of contribution to our constructive exchange, let me expand on one point, included in the quotation above, namely the concept of cooperation.

Cooperation is different from *collaboration*, which is different from *coordination*. Agents coordinate when they simply do not hinder each other while going their own ways. Imagine Alice and Bob cooking and eating their own meals when they want and as they wish, using the same kitchen. They coordinate their actions as long as neither of them represents an impediment for the other. Less metaphorically, when markets work correctly, they are good at creating coordination, e.g., through competition between Alice and Bob. Collaboration requires coordination, but it also includes sharing tasks: Alice may contribute the appetisers and the drinks and Bob the main course, in our example. Markets are less good at creating relations of collaboration in the absence of incentives. Cooperation needs even more, for it implies sharing the whole process: Alice and Bob do the shopping and the cooking together. They co-design, co-create, and co-own the meal, so to speak. Markets do not perform well when it comes to cooperation unless the law intervenes. This is where I find my own work on

infraethics and the order ethics approach complementarily helpful. For global problems require more than coordination or even collaboration, they require cooperation: a sharing of decision processes, choices, and implementations of policies that touch the lives of millions and sometimes billions of people. We only need to think about the pandemic or climate change. So, markets are necessary mechanisms, but they are largely insufficient without political will and normative incentives. [388]

Let me close with a couple of clarifications. I may be wrong in my analysis, but when I argue that we need to upgrade our ontology, what I mean is that I would welcome a relational way of thinking and conceptualising the world, including above all socio-political issues, as mainstream as opposed to an intellectual effort that has tried to make a difference since Plato. Far from me to say that we never reasoned relationally, or that there are no important precedents in understanding the world relationally. This would be a mistake too silly to make. What I am arguing is that, if we look at how we frame contemporary issues, we still see a Lego-like approach being the default approach. Referring to the debate about AI touched upon by the comment, for example, how many times do we still hear that it is a 2 or 3 players game, US, China and maybe the EU? This is what I mean when I say that we should change our perspective.

Finally, I agree that the second half of the twentieth century reacted to the horrors of the first half by implementing a meta-project that would not offer a social project but only the protection of individual projects. This has been a significant development and I hope a point of no return, at least for liberal democracies. However, today, we also need to find a better middle-ground. We need to ensure that Alice and Bob can pursue their own projects, but also help them to have a project as a couple, to use my previous analogy. Because the global problems we are facing can only be solved together, cooperating, not just individually and merely coordinating. To use a recent example, it is only if the G7 and then the G20 cooperate that the problem of tax abuses by multinationals and online technology companies can be tackled. Even the whole EU would be insufficient, if working alone. The American constitution begins with »We the people…« and it is precisely that »we approach« that I am defending in the article, not as an alternative to, but as a necessary complement of the individualism to which we are so accustomed: we must walk on two legs, have protection of individual projects and promotion of social projects. We

need coordination, but also collaboration and cooperation. This is how I would analyse an »institutional understanding of ethics«, an important remark that concludes their interesting comment.

Reply to Catrin Misselhorn

I expected a comment about the article in this collection, but the text by Misselhorn is about a book I published in 2019: *The Logic of Information – A Theory of Philosophy as Conceptual Design* (Floridi 2019). Putting aside the surprise, I begun reading it as a review of the book, but, actually, it is only a series of objections to my replies to four potential objections that I imagine one may formulate (both conceptually and historically) to my interpretation of philosophical questions as open questions, that is as questions that remain open to reasonable disagreement even when the parties involved have all the factual, scientific, logico-mathematical information one may wish them to enjoy. If the readers have already lost any interest, I fully sympathise. However, if they are still reading, then, regrettably, I must confess that I have not learned anything from the objections. Of course, this is my [389] problem and my loss. But I have a justification. The objections appear to me so unrelated to what I mean, state, and argue that, conceptually speaking, they are not even wrong. Instead, they remind me of the famous remark by Pauli: »*Das ist nicht nur nicht richtig; es ist nicht einmal falsch!*«. Let me give you one example. I hope everybody will agree that »what is the result of 1+1?« is a mathematical question, even if extremely simple, perhaps too elementary to bother anyone who is not a child (but mind that it takes hundreds of pages to prove that the answer is 2 in *Principia Mathematica*). Mathematical questions studied by mathematicians are way more complicated and more consequential. Yet, this takes nothing away from the mathematical nature of the question about the sum indicated above. Mathematical questions can be that simple and elementary. So, when someone objects that a philosophical question cannot be an open question because some open questions are too simple and elementary to qualify as philosophical, like »should I wear my hair shorter?«, the reply is similar: that is still a philosophical question, just one that is not very interesting and consequential. There is not even a bullet to bite; this is just plain common sense. I am sure there is a stage in life when it is a crucial, significant question

for someone. But it is not just the sort of foundational, consequential question addressed by philosophers, who may start questioning what the alternatives are (long, very long, short, very short, etc.), and whether some of them are dictated by social, or peer pressure or perhaps fashion or maybe health, etc. to finally get to something that is richer in significance and consequences and deeper in insights. But philosophical questions too, can be that simple and elementary. Just check those asked by children.

Things do not improve as the text progresses, and the objections end with a rather odd description of the method of abstraction – something quite ordinary in Computer Science where it is studied in the context of Formals Methods – which I could not recognise, and indeed quite distant from the (textbook) material presented in the book. I won't bother the reader with all this; I only wish to stress my inability to follow Misselhorn's text. I have not recognised any of my ideas in the comment. The last part is particularly baffling. Here is an example:

> Floridi's new political ontology tends to obscure these dangers. The replacement of the individual as the normative foundation of society by a relational view that reduces it to a node in a functional system lends itself readily to technological solutionism which goes against the spirit of liberal democracy. Discarding the idea of the free and equal moral person as the normative basis of political theory is tantamount to affirming the practices of the Tech Giants even if Floridi wants to give them a positive spin with infraethics. (154)

Nothing could be more distant from what I wrote and argued. I find it reassuring that several other commentators in this collection have understood the points I have sought to make and criticised them with insightful clarity. It shows that it is doable. Of course, readers have the right to misread authors and misinterpret their intentions as they wish. Sometimes it is even helpful for the development of their own ideas. But authors have the right to be astonished by such a lack of understanding, refuse to have words put in their mouth, as the saying goes, and not engage with something they never wrote or meant in the first place. [390]

Reply to Malte Rehbein

The comment by Rehbein is interesting because I have learnt from it (as from some, though we have seen not all, comments in this collection) where I need to be clearer and more explicit. I am saying this because I believe I would subscribe to almost anything he writes (more on the »anything« qualification presently). However, in the future book, I shall still resist the temptation of moving from discussing problems and solutions to discussing people and their theories. If I am wrong, I won't become right just because I make such a shift; if I am right, the shift may always follow later (and if anyone is interested in doing such interpretative work, I shall be most grateful and honoured). For now, *ars long vita brevis*, as they used to say, life is too short, and I am keen on exploring the ideas discussed by Rehbein in his comment, not people. So, I would rather run the risk of reinventing this or that wheel than spending a lifetime wondering whether the wheels I need have already been invented, by whom, and why, and whether they are really like the ones I need or just similar, which ones work better, and so forth. As Montaigne once wrote (I go by memory, I hope it is Montaigne), one cannot do research (I think he says explore) without losing sight of the coast. So here I am, lost. This is not a license to be lazy though; therefore, in terms of my contribution to this asynchronous dialogue with Rehbein, let me take advantage of a very helpful paragraph, on p. 157, to clarify some of my thoughts. I will structure the paragraph into a conversation between R (Rehbein) and (F):

R: »I would like to argue that the world is not secular.«

F: I agree, but I would like to add: unfortunately. A secular world has a better chance of being more tolerant than a non-secular one (by the way, I am not an atheist, I am a religion-friendly agnostic, and I would give anything to reacquire my faith, I just seem to be unable to get it back no matter how hard I try).

R: »It is not binary, neither ontologically nor in terms of an information divide.«

F: I agree, but in the same sense then it is not analogue either; on this, I agree with Kant, discrete vs continuous are ways in which we conceptualise reality (Floridi 2009).

R: »Technology is not the only solution, but part of the problem and it should be treated as such.«

F: This is imprecise. Technology can be part of the solution – do we really need to stress the immense benefits of technology at all levels? I am reminded of it every time I go to the dentist – but can also be part of the problem – ditto – so it is up to us to make sure that only the first half takes place.

R: »Capitalism is not a compelling prerequisite (markets are).«

F: I am not sure what »compelling prerequisite« means. However, if it means that we can solve our environmental and social problems by getting rid of capitalism, or by-passing it, or stopping it, etc., then I wish that were true, but I fear we better be realistic and harness capitalism and its energies to solve the problems we have. This is why politics, legislation, and governance are so crucial.

R: »The human condition, together with a new contractual definition of global [391] equality and justice, well-being and welfare beyond materiality and consumption should indeed be the starting point of any human project.«

F: I agree. I would add that we need to create enough wealth for the billions of people who live so miserably, though. »Beyond materiality« is fine as long as it is not the kind of materiality that determines the availability of food and shelter, decent living standards, human rights, jobs, health care, safety, etc. Call that »good materiality«, and we are on the same page. I intensely dislike consumerism, but we need more »good materiality« for billions of people.

R: »However, information is a necessity, but not a sufficiency to serve as a core concept for a new ethical and political framework.«

F: If I understand this correctly, I agree, and strongly doubt anybody could disagree.

R: »What is required might not be a new ontology, but a new inter-generational social and environmental contract.« – End of the paragraph (p. 157).

F: In the book, I have suggested replacing the social contract with a sense of ontic trust. Let me close this reply and the whole set by

summarising what I mean by it (the reader interested in knowing more can check chapter 15 of (Floridi 2013)

A straightforward way of clarifying the concept of ontic trust is by drawing an analogy with the idea of »social contract«. 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 an actual, implicit, or merely hypothetical agreement between the parties (e.g., the people and the sovereign, the members of a community, or the individual and the state) constituting a society. 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 artificial agent created to enforce 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, individual, informed 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, that all human (I shall drop this specification henceforth, unless this generates confusion) agents cannot but sign when they come into existence, and that is constantly renewed in successive generations.

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 some specific 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, [392] 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 world, including all existing agents and patients;
- 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.

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, 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 an agent may break the ontic bond 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 recognising 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, Alice 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 moral agents' career as strangers to the world; we should end it as friends of the world.

Bibliography

Floridi, L. (2002), *Sextus Empiricus: the transmission and recovery of pyrrhonism*, Oxford.
– (2008), »A Defence of Informational Structural Realism,« in: *Synthese* 161(2), 219–253.

– (2009), »Against digital ontology,« in: *Synthese* 168(1), 151–178.
– (2013), *The Ethics of Information*, Oxford.
– (2014), *The Fourth Revolution – How the Infosphere is Reshaping Human Reality*, Oxford.
– (2015), »Toleration and the Design of Norms,« in: *Science and Engineering Ethics* 21(5), 1095–1123.
– (2016a), »Hyperhistory, the emergence of the MASs, and the design of infraethics,« in: *Information, Freedom and Property: The Philosophy of Law Meets the Philosophy of Technology* (ed. Hildebrandt; van den Berg), Chapter 7, London.
– (2016b), »Mature information societies – A matter of expectations,« in: *Philosophy & Technology* 29(1), 1–4.
– (2016c), »The method of abstraction,« in: *The Routledge Handbook of Philosophy of Information*, London, 66–72. [393]
– (2016d), »Tolerant Paternalism: Pro-ethical Design as a Resolution of the Dilemma of Toleration,« in: *Science and Engineering Ethics* 22(6), 1669–1688.
– (2017), »Infraethics – on the Conditions of Possibility of Morality,« in: *Philosophy & Technology* 30 (4), 391–394.
– (2018), »Semantic Capital: Its Nature, Value, and Curation,« in: *Philosophy & Technology* 31(4), 481–497.
– (2019), *The Logic of Information: A Theory of Philosophy as Conceptual Design*, Oxford.
– (2020a), »The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU,« in: *Philosophy & Technology* 33(3), 369–378.
– (2020b), *Il Verde e Il Blu – Idee ingenue per migliorare la politica*, Milan.
Roberts, H./Cowls, J./Morley, J./Taddeo, M./Wang, V./Floridi, L. (2021), »The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation,« in: *AI & SOCIETY* 36(1), 59–77.