NOVATION

Critical Studies of Innovation

cover image: David Cristian

NOVATION

Critical Studies of Innovation

[Online Journal]

Second Issue
2020

Responsible Innovation (RI) in the midst of an innovation crisis

Guest Editors

Lucien von Schomberg, University of Greenwich

Vincent Blok, Wageningen University

Hosted by *l'Institut national de la recherche scientifique*, Centre | *Urbanisation Culture Société*. Montreal, Canada

About Us

The international journal *NOvation: Critical Studies of Innovation* was launched to contribute to the rethinking and debunking of innovation narratives in STS (Science, Technology and Society) and STI (Science, Technology, and Innovation). There is a need to critically examine studies of innovation and obtain a clearer portrait of innovation than the depiction this field has been accustomed to. The journal questions the current narratives of innovation and offers a forum for discussion of some different interpretations of innovation, not only its virtues, but also its implications. In this sense, NO refers to non-innovative behaviors, which are as important to our societies as innovation is. Failures, imitation and negative effects of innovation, to take just some examples of non-innovation or *NOvation*, are scarcely considered and rarely form part of theories of innovation.

ISSN 2562-7147

Contact Us

novation@inrs.ca

Table of Contents

Editorial Presentation: Responsible Innovation (RI) in the midst of an innovation crisis

Lucien von Schomberg* and Vincent Blok**

** University of Greenwich*

*** Wageningen University*

The concept of Responsible Innovation (RI) occupies a central place in the discourse on science and technology, especially in the context of the European Union (EU) but also within academia. This concept is guided by the idea of steering science and technology towards societally desirable outcomes, particularly in response to normative objectives such as Sustainable Development Goals (von Schomberg, 2019). Visions of RI typically propose that to innovate responsibly requires a permanent commitment to be anticipatory, reflective, inclusively deliberative, and responsive (Owen *et al.*, 2012). They also emphasize the need for open access, gender equality, science education, ethical standard in conducting experiments, and democratic governance (European Commission, 2020).

However, the societal purpose of RI fundamentally conflicts with the imperative of maximizing economic growth inherent in today's innovation climate (von Schomberg, 2022). This conflict points to a crisis in which innovation struggles to serve public interests insofar private interests continue to be prioritized. The magnitude of this crisis is also reflected within the RI literature itself, where the political ambition to exceed the privatization wave is summoned to a techno-economic concept of innovation (von Schomberg & Blok, 2019). This issue of *NOvation – Critical Studies of Innovation* brings into question to what extent innovation necessarily relates to the market, whether it is possible to develop an alternative concept of innovation that is separated from economic ends, and how we can conceptualize, for example, a political understanding of innovation. What really is innovation? While all seven contributions share the aspiration to critically reflect on these questions, they each offer a distinct and original perspective in discussing the relation between innovation, technology, politics, economics, and responsibility.

In the first research article, Bedreddine (2022) draws on the interdependency of politics and economics to examine the emerging context of RI in France. In doing so, RI is shown to be constituted in a space where agents of the private sphere and the public sphere converge. Through empirically analysing interactions that take place in this space, including those between innovation managers, politicians, executive directors, and the wider public, the article investigates the way in which innovation transforms the fields of economics and politics in France, resulting in a loss of autonomy for both.

In response to the changing nature of innovation in the digital age, the second research article invites the RI discourse to revisit their foundational narrative (Bryce *et al.*, 2022). It explores to what extent RI is anchored in underlying assumptions about contemporary technologies and, in turn, what limitations this faces in today's increasingly digital context. As such, the authors aim to broaden the horizons of RI, highlighting that the potential to steer innovation towards societally desirable outcomes depends on the awareness researchers and practitioners have of digital technologies and so-called metatechnologies.

The third research article critically accounts for how mid-stream actors deal with tensions between a commitment to RI and anticipated market requirements (Frost *et al.*, 2022). Through conducting exercises that build on Socio-Technical Integration Research (STIR), the authors point to "the underlying assumption that marketability of prospective outcomes is not one objective amongst others but the precondition for all others". Social and environmental values are only considered insofar they are adopted by a techno-economic paradigm of innovation. To this end, the article calls for greater efforts beyond midstream constellations to contest the resilience of the techno-economic paradigm of innovation.

The concept of innovation lacks a strong conceptual understanding both within and beyond the RI literature. To this end, Michels (2022) argues that "innovation is inescapably normative" and proposes a new definition in which innovation is understood as "*ethical change that delivers substantial applied value to beneficiaries of a domain*" (original emphasis). Through articulating this novel definition, the fourth research article rethinks the relationship between innovation, technology, and the marketplace, ultimately refining the meaning of RI.

While Michels (2022) points to the normativity of innovation, Penttilä (2022) argues that the operationalisation of such normativity requires a strong political dimension. Particularly in response to the phenomenon of depoliticization, structurally underpinned by economic incentive, the fifth research article urges frameworks of RI to "adopt a *political* conception of responsibility in order to safeguard the legitimacy of the values and outcomes it deems societally desirable" (original

emphasis). Drawing on the work of Hannah Arendt, it accounts for the interrelation between responsibility and politics, and in doing so, contributes to the politicization of RI.

The interrelation between responsibility and politics is further reflected in the RI dimension of anticipation, as denoted by the sixth research article. Here, Rodríguez *et al.* (2022) argue that the scope of anticipatory practices is specified by the socio-political context in which they take place. In the EU research and innovation policy context, they identify such practices with a "disruptive-limiting" duality. On the one hand, the emergence of RI aims to facilitate a critical and radically open debate on the underlying purposes of innovation systems. On the other hand, the dominant techno-economic imperative limits such as debate to "normative milestones that are prefixed and impervious to debate".

In the final research article, Tabarés (2022) employs a RI perspective to critically assess the development and challenges of Open Access (OA). While OA provides several opportunities to transform the landscape of academic publishing, under the sway of digitalization it has "reinforced the oligopoly of for-profit academic publishers". To this end, the article argues that OA should not exclusively focus on making scientific articles widely available, but more fundamentally, contest the exploitation that takes place in the growing "platformization" of academic publishing.

REFERENCES

European Commission (2020). *Horizon 2020: work programme 2018-2020: Science with and for Society.* Resource document. https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-swfs_en.pdf

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: from Science in Society to Science for Society, with Society. *Science and Public Policy,* 39(6), 751-760. https://doi.org/10.1093/scipol/scs093

Von Schomberg, L. (2022). *Raising the Sail of Innovation: Philosophical Explorations on Responsible Innovation*. Wageningen University. https://research.wur.nl/en/publications/raising-the-sail-of-innovation-philosophical-explorations-on-resp

Von Schomberg, L., & Blok, V. (2019). Technology in the Age of Innovation: Responsible Innovation as a New Subdomain Within the Philosophy of Technology. *Philosophy & Technology*, 34, 309-323. https://doi.org/10.1007/s13347-019-00386-3

Von Schomberg, R., & Hankins, J. (2019). *International Handbook on Responsible Innovation: A Global Resource*. Edward Elgar Publishing.

Responsible innovation in France. A proxy allowing agents of the political and economic fields to interact[1]

Samir Bedreddine*

* *Paris Saclay University*

ABSTRACT

This article examines the issue of responsible innovation in France. To describe its underlying mechanisms and logics, we retrace the advent of innovation in public policy and its reception in French firms, showing how responsible innovation has become the cornerstone of public-private interactions. The legal and administrative context of innovation in France, on the one hand, the emerging departments and managers of innovation in the large groups, on the other hand, participate in producing spaces where agents of the political and economic fields converge. Such situated interactions hinge on shared world views, values and tools. Innovation managers, executive directors of large firms, some French politicians and public servants seize upon responsible innovation and create areas regulated by specific norms and values. In these shared spaces, responsible innovation is the star object, the proxy for exchanges between the economic and the political fields.

Keywords: Innovation; Responsibility; Politics; Management; Research; Large Firms.

Proposal Submitted 21 August 2020. Article Received 9 April 2021. Reviews Delivered 17 May 2021. Revised 26 June 2021. Accepted 6 September 2021. Available online 28 February 2022.

[1] Translation by Gabrielle Varro.

INTRODUCTION

For approximately ten years now, innovation has been presented as the summum of French economics. The apparently sudden interest of the State is in fact part of a long political history, which gathered momentum in the 1960s, punctuated by several interventionist laws, whose ambition it was, and still is, to produce a "national system of innovation" (Freeman, 1995). The period bears witness to the "social construction of the public problem" of innovation (Bourdieu, 2012, p. 53). First conceived as *innovation through research*, then as *innovation by research*, the present-day situation postulates *innovation beyond research*. The French State aims to organize relations between Industry, Government and Academia, a configuration that some scholars have baptized the Triple Helix (Etzkowitz & Zhou, 2017).

In such a context, politics and economics are closely related (Bourdieu, 2005). Moreover, the accusation of "hyper-interventionism" (Lebaron, 2016) seems not to apply to the private initiatives that accompany innovation. The Public Investment Bank (also known as *Bpifrance*), as well as research tax credit, happen to be the primary financial resource for private innovation in France. As a result, in what way do the itineraries, practices and values of the agents of innovation transform interactions between politics and economics? Responsible innovation, defined here for the French context, emerges as the cornerstone of relations between the public authorities and the organizations. The very same responsible innovation is benefitted by the publicity the public authorities give it; it infiltrates the organizations thanks to their directors' determination, as well as to the importance of the personal itineraries of those I call innovation managers and to their positions on the question of values. Also, responsible innovation is part of the interactions observed between the public and the private sectors.

In this article, we scrutinize those transformations at the level of both private companies and public research institutions, applying Pierre Bourdieu's field theory. Studying responsible innovation in the light of this concept allows theoretically reconstructing social spaces regulated by laws where individuals are more or less able to participate in the battles induced by the agonistic nature of the field. Specific capitals and regulation permit agents to accumulate and use a certain number of resources that make them legitimate. Symbolic and cultural capital allows them to conquer spaces for self-expression which the logics of innovation then disrupt. We show how responsible innovation is pivotal for interactions to develop between the political and the economic fields.

In Part One, we present the laws that have structured innovation in France since the 1960s, illustrating the strong implication of the State. Our ambition here is to

demonstrate the strong implication of the State in the social configuration of the economy (Bourdieu, 2005; Etzkowitz, 2002, p. 139-144), a State striving to build a field of possibles for those individuals who want to become professionals (Chiapello & Gitiaux, 2009; Maitre & Bourdieu, 1994). This section reveals how this develops over the long term and allows observing how the notion of responsibility (Jonas, 2013) nourishes the Triple Helix configuration Government / Industry / Universities. This, in turn, is a chance to see how renewing political personnel (Michon, 2019) reinforces pre-existing dynamics.

Part Two allows us to broach the ways political injunctions were received in the private sector. We will show how industrialists had to cope with both political injunctions and organizational obligations. Private business at the time had to deal with political considerations and individual acts, while itineraries, values and ethics facilitated the emergence of the notion of responsible innovation.

The last part will broach the practices common to both the political and the economic fields, and the places where they meet. We explain how responsible innovation can be considered a proxy for their interactions. In fact, the State counts on private companies to let it pursue its ambitions for the energy transition as well as for other questions that can also interest the organizations (Bereni & Prud'Homme, 2019). These zones of interaction are produced and activated by a collaboration between managers of innovation, public servants and the top authorities in each field.

This article is based on a qualitative survey (interviews and observation). To study their activities, their itineraries and the positions they defend, we met with the innovation managers of three large French companies of public works, telecommunications and energy. Personnel of public and political institutions in charge of innovation were interviewed in order to document their habitual practices. We also analyzed legal texts and records of parliamentary hearings bearing on questions of research, innovation and industry. Our data also includes observation carried out in fairs, prize award ceremonies and areas dedicated to innovation in the large firms.

POLITICS OF RESEARCH AND INNOVATION IN FRANCE

Innovation through research

During the 1960s and 1990s, the "social construction of the public problem" of research and innovation (Bourdieu, 2012, p. 53) grew out of the creation of public commissions (Bourdieu, 2012, p. 47). Elected officials preferred to speak of basic research, applied research, and development research rather than of innovation. On November 16, 1966, the law proposing to "create public research organizations" came

before the National Assembly. On November 30, 1966, Paul Thillard, reporter, justified the creation of a certain number of public institutions. The verdict was indisputable and unanimous: France was sorely lagging behind her competitors. Mainly due to the progress of science (Gingras & Villedieu, 2010, p. 130), the Great War (1914-1918) had amputated a significant number of the dynamic members of French science: "the second world war plunged our country's science into a deep sleep, while abroad it flourished, producing the radar, rockets, jet engines, electronics, atomic energy, antibiotics".[2] The opposition exacerbated the sense of urgency; François Mitterrand compared France's skills in computer science to the competition between the United States and Europe. In a bipolar world focused on nuclear power and the Space conquest (Joly, 2017; Edwards & Hecht, 2005) the idea of a "techno-nationalism" (Edgerton, 2013, p. 150-158) *à la française* became all-pervasive in the political field.

To live up to that challenge, it was first of all necessary to reinforce the protection of intellectual property, by improving the law on patents and licenses. Protecting the fruit of research would not slow up innovation (Amable *et al*., 2006) but would prevent a "brain drain" and a loss of results. It would also call for tightening the links between public research and the private sector. Creating[3] the National Agency for the Commercialization of Research (*Agence Nationale pour la Valorisation de la Recherche,* ANVAR), a major part of the projected Law, therefore meant organizing and supporting relations between Academia and economics. Though passing from public research to development and applied research was unanimously considered inefficient, there was much disagreement about how to avoid the attempt of industry to cow, or even dominate (Thébaud-Mony, 2014) scientific research (particularly fundamental research) in the public sector and guarantee the autonomy of public researchers. The profession of research scholar decidedly called for a transformation, and the figure of the new "man of science" (Shapin, 2006), actor of a "science necessarily allied to industry" (Lamy, 2020, p. 23), seemed to take root in the political field. Alain Peyrefitte consequently seemed to be echoing a phrase of General de Gaulle's,[4] apocryphal but telling of the spirit of the time: "The onus to search – and above all to find – falls upon nations as it does upon industry, for fear of being irreversibly passed by."

The law "creating public research organizations" was published in the *Journal officiel* of January 3, 1967. It was the time of *innovation by research*, or the superiority of scientific work over innovation. On November 30, 1966, the debates on the law

[2] Speech by Paul Thillard.

[3] This law was debated in 1966 and adopted on January 3, 1967.

[4] Charles de Gaulle's phrase was wittier: «You can find researchers who search, but you must search for researchers who find. «Des chercheurs qui cherchent, on en trouve. Mais des chercheurs qui trouvent, on en cherche».

proposal contained 375 occurrences of the word "research" and 105 of the term "scientific", whereas the word "innovation" appeared only four times. This lexical measure shows up a salient fact of the law and the debates, namely that, in 1967, the issue was not to separate scientific research from innovation, but to stress that the latter was part and parcel of scientific work.

The Law of 1967 stated the intentions of the public authorities in terms of research. In 1972, the legal frame for investing in innovation was defined. The Law of 1972 created the "financial company for innovation" (*société financière d'innovation*) and provided for the prerogatives of the State in the companies, their organization, functioning and taxation. In 1979, a decree stipulated that the mission of ANVAR was "to enhance the results of scientific and technical research and to promote innovation and technological progress" – one of the first occurrences of the term innovation. The laws voted on July 15, 1982 and December 23, 1985 followed up that initiative by making innovation one of the prerogatives of public research and teaching. What was at stake for the public authorities was to produce agents capable of making the "national system of innovation" work:

> Schools, universities, and continuing education at all levels, as well as the public services of radio and television must all promote the spirit of research, innovation and creativity and participate in the development and transmission of the scientific and technical culture. (Article 7, 1982)

Innovation beyond research

Between 1960 and 1990, the public service of teaching and research, and the administration of economic policies, took innovation in hand. The fact there was a considerable amount of State intervention confirmed that innovation was a direct consequence of research. In the years following, fluctuating representations became part of a continuum, and scientific discoveries needed to fit into an economy hungry for new techniques. This made it necessary to pass a law on the status of civil servants in public research.

In the late 1990s, Prime Minister Lionel Jospin returned to the question of innovation and reaffirmed the commitment of the State. The main point of his speech on July 24, 1997, was to promote and accompany relations between the public and private sectors. He insisted that knowledge must circulate outside Academia, and suggested public researchers be encouraged to create their own businesses: "In the same vein, the Government intends to take significant action so that scholars who so desire be enabled to create a firm to commercialize the fruit of their research and benefit from public and private funds not available today". It was the start of *innovation beyond research*. On May 10, 1998, the prime minister concluded the "Foundation for Innovation" conference by exposing his vision of an effective French system of

innovation.[5] The State must guarantee the circulation of knowledge and back private initiatives. The importance of innovation spurred the prime minister to endorse increased State interventionism: "I believe on the contrary that strong public intervention is justified in a sphere where the benefits for the community exceeds private interests. [The State] must also guarantee that innovation and growth do not endanger social cohesion and that everyone benefits. It must remain the guardian of national cohesion". The power of the State must accompany innovation, ensure it is "responsible", and therefore dispose of the right to exercise an "ethical" overview of the innovations produced by the private sector. The law of July 12, 1999, on innovation and research authorizes civil service researchers to create businesses and stipulates how roles and capital are to be shared out in such companies.

Between 2000 and 2010, several dispositions and systems were deployed to support innovation in the economic field. An "economic pole" emerged in "political itineraries", and "hyper-interventionism" was omnipresent despite being branded the "*mal français*" (Lebaron, 2016). In fact, public reports (Née *et al.*, 2017) called on the powers that be to encourage private investment in Research and Development by creating a legal and financial framework. The Law on Finances for 2004 generated the status of "young innovative enterprise" and defined the criteria for obtaining the label and its correlative advantages, particularly concerning taxation. In a ruling of June 29, 2005, the public establishment OSEO[6] replaced ANVAR. The new public organization was supposed to "promote and support innovation, particularly in technology, and contribute to the transfer of technologies", as well as "encourage the creation, development and financing of small and middle-sized firms". In 2012, the Public Investment Bank (BPI) replaced OSEO, once and for all institutionalizing a socially and ecologically responsible innovation. Via the BPI, companies were to be led towards "responsibility", to sustain "durable growth, employment and the strength of the economy" by participating in the "development of sectors of the future, of digital conversion, and of social and solidarity-based economies" and by "supporting the implementation of the energy and ecology transition". The BPI was oriented "in priority towards the Very Small Firms, the Small and Middle-sized Firms and the firms of intermediate size, particularly in the industrial sector". At the start of 2014, the French section of the "Horizon2020" program set up a European public action in favor of research and innovation.[7]

[5] Lionel Jospin's speech in 1998 leans largely on the «Rapport de mission sur la technologie et l'innovation» submitted by Henri Guillaume in March 1998.

[6] https://www.bpifrance.fr/nos-actualites/oseo-filiale-de-la-banque-publique-dinvestissement-lance-un-nouveau-fonds-de-garantie-pour-soutenir-la-tresorerie-des-pme-et-des-tpe.

[7] https://www.horizon2020.gouv.fr/cid75845/lancement-programme-horizon-2020-decembre-2013-dossier-presse.html.

In May 2017, Emmanuel Macron was elected President of France and redrew the political landscape. There was a "massive arrival of political neophytes", whose social attributes and political itineraries were far from those previously noted (Bargel, 2014; Dolez *et al.*, 2019, p. 220). The 2017-2022 mandate of the Presidential Party – "*La République en Marche*" – brings together individuals endowed with particular socio-demographic characteristics. Arriving mainly from the private sector (64% of the Party's representatives), they make up an "economic elite" of "entrepreneurial good will" (Dolez *et al.*, 2019, p. 224). In an interview, Amélie de Montchalin, elected in June 2017, presents her itinerary and her work as follows:

> I was trained as an economist, I have a Master's from HEC[8] specialized in economics, a licence (B.A.) in applied economics from Dauphine (Paris University). I resumed my studies at the Kennedy School of Harvard, where I did my Master's in Public Administration, doing a lot of economics and thinking about the reforms. Ten years in two firms, one of which was a branch of a large French bank, where I was the economist in charge of the Euro zone … then with a big French insurance company, where I was in charge of foresight and mid-term strategy. All in all, for three years I worked for the ComEx, i.e. the worldwide executive committee on the risks to insure in the future. […] Climate change and also understanding the distribution of what the States and what the private actors will do. Therefore, in connection with the European commission, the G20 and the U.N.

In 2019, Amélie de Montchalin was appointed minister. Her itinerary is an extreme case that shows how the private sector has infused the field of governance as well as personal profiles, dominated by economics. Political women and men now discuss foresight, the future, and risk management. The language they use includes the vocabulary of innovation introduced by the representatives of innovation (Bedreddine & Noûs, 2021), already prevalent in the large firms.

Also, in May 2019, a law bearing on the growth and transformation of companies (Loi PACTE) was voted in. Section 2 of Chapter III aims to "reconsider the place of a firm in society". The law provides for the creation and conferring of "labels of corporate social responsibility", based on various criteria, by introducing a variety of legal and administrative dispositions that enhance companies' social and environmental commitments. "A company's statutes may highlight its *raison d'être*, i.e. the principles it means to observe and to which it may allocate resources in the course of its activities". Thereby "a company can publicly declare it has a mission, as long as the following conditions are satisfied: 1) Its statutes specify a *raison d'être*, in the sense of Article 1835 of the Civil Code; 2) Its statutes specify one or several social and environmental objectives, that it considers its duty to follow up in the course of its activity; […]". The firm is called upon to broadcast its "social worth, by taking the social and environmental stakes of its activity into consideration". The public

[8] Haute École de Commerce, one of the elitist French grandes écoles.

authorities are bent on setting up and reporting the efficacy of the legal system by enhancing the "behavior and strategies that correspond to a public charter of good practices by acknowledging the extra-financial performances of the firm […]".

Amélie de Montchalin was very involved in writing up that law:

> I really like the idea behind the PACTE law, the idea that companies have a mission to perform […] and every time we scrutinize that mission in a context that's changing in a changing world, well, the fruit of that scrutiny is often innovation.

The Minister went on to add that innovation can only be responsible and that she would like to impose new norms:

> Well, innovation – if I go back to my three points – must be aware of its consequences. Today, innovation that leads to consuming even more resources, or that puts even more gas into the atmosphere, or that creates inequality, should not be implemented, because it goes against challenges we already have to face… so, from a normative point of view, people must be consequential.

Private actors agree to that political definition of responsible innovation, while placing the State at the hub of its economic mechanisms. Firms and their representatives use the definition to develop new discourses and new practices.

FIRMS AND THE POLITICAL RHETORIC ON RESPONSIBLE INNOVATION

Attentive executives

> The role of a firm goes way beyond profit-making. It must also play a strong social role, be committed to its stake-holders (employees, clients, providers, shareholders, States, etc.). Its leaders and employees are duty-bound to make sense of their firm's activity.[9]

The above post, published on LinkedIn on September 8, 2020, by the general director and president of the Thalès company, demonstrates a fair degree of consensus with the political norms of corporate responsibility. The same principles seem to be voiced in a report by the Haut Comité de gouvernement d'entreprise, which claims as its own the legal dispositions of the PACTE law (2020, p. 17). One of the report's proposals consists in partly linking executive managers' dividends to the sustainability and durability of their firm's activity[10]. The idea underlying this sort of proposal is to make

[9] « La raison d'être, une boussole précieuse au cœur de la crise », https://www.linkedin.com/pulse/la-raison-d%25C3%25AAtre-une-boussole-pr%25C3%25A9cieuse-au-c%25C5%2593ur-de-crise-patrice-caine/?trackingId=JBxULlWeQsa0atz0hlLR7w%3D%3D

[10] The report of the High Committee that brings together two associations of leaders of industry explains: « It is no longer acceptable that a leader's variable compensation not include environmental criteria. The High Committee expects RSE

the parameter of responsibility part of the firm's strategy, in particular by a trade-off of RRI, based on identifying and dealing with the social and environmental consequences – negative as well as positive – throughout the innovation project (Paredes-Frigolett, 2016).

During the 2010 decade, talk of a new conception of innovation began to circulate in the firms, while a number of reports and public declarations came to light. The agents of social change (Rogers, 1995, p. 335) began the job of producing and perpetuating the belief in innovation. In 2014, the future director of innovation at BatiCorp E[11] began discussions with the leader of the group:

> The decision to create a department of innovation emerged from the discussions I had with the general director four years ago. The digital transformation of the skills at BatiCorp E was striking. That transformation, associated with the energy transition, led first to questioning the "smart" model […]. That's when the decision to put a single person in charge of everything connected to smart in general was made. My work with the general director led to imagining a department of innovation that would allow bringing everything together, to deal with the subject transversally and make headway in the field of energy transition…[12]

A campaign was launched to convince people of the need to create a department of innovation and a generational effect became apparent, when younger people entered the fray:

> It must be said that those who convinced the ComEx to do it were the young managers. […] At BatiCorp there are forty young guys, the top 10% managers, who are between 25 and 40 and who got together and said careful – to make it short – we've got to have a sort of innovation cell…[13]

Young managers with a particular vision of the firm and its business were called in to create the group's innovation system of norms and values (Granovetter, 2017). Giving a meaning to employees' activity became one of the missions of responsible innovation, which it accomplished by "attracting a sponsor", thus, in fact, establishing innovation as a vertical, hierarchical privilege:

> Innovation must make sense… the company itself must rediscover its *raison d'être* thanks to innovation, particularly in the large groups. That being said, it's the executive director's place to explain why, the innovation director's role to list the

criteria to be defined precisely, clearly, pertinently and in such a way as to include the social and environmental stakes of the firm. Simply referring to an application of RSE (Corporate Responsibility) criteria or to an in-house RSE program or to what's at stake generally, without defining them clearly, is not sufficient ».

[11] The names of companies have been changed.

[12] Interview carried out on 03/04/2018 with Nadège A. (age 55), at the time Innovation Director for the BatiCorp E group, who holds a University diploma and a degree from a business school, and is specialized in finance and financial engineering.

[13] Interview carried out on 25/01/2019 with Wilfried C. (33), energy engineer by training, then director of acceleration and entrepreneuriat in the department of innovation of BatiCorp. A graduate of a school of engineering in energy (Centrale, Nantes), he did his final year in a double course in market finance. He holds a degree in applied mathematics at the School of Mines, where, after a Master's at Stanford (advanced management program), he is continuing a thesis begun in the U.S.

> possibilities, and the executive director's to make his/her choice among the strategic foresight scenarios under study.[14]

Employees working in innovation departments may find themselves caught up in a move to reclassify the workforce (Chiapello & Gitiaux, 2009). It is also an opportunity to take on new workers, fresh out of business schools or of university "innovation curricula".

Studies on generation Y, also described by the dubious categorization of *Millenials* (Bennett *et al.*, 2008; Negroponte, 1996; Ughetto, 2018, p. 163), or of generation Z, also incite the upper strata of the organizations to "hold on to their talents". "Fifteen years ago, people stayed approximately six years at BatiCorp. Today it's one year and 8 months".[15] It is necessary to retain but also to attract "talents": "The ComEx said yet another thing: that we have a lot of difficulty attracting talents at BatiCorp, it doesn't have a great reputation". Innovation becomes a form of internal and external communications aimed at present and future employees:

> The fact there is an overall impetus due to BatiCorp E's position as leader in innovation on the outside also reflects on the inside… and people identify more and more with the fact of being a leader accompanying the digital transformation and its clients' energy and ecological transition…[16]

The approach attracts employees registered in programs for human resource management (Cihuelo, 2020) or in new, so-called innovation activities, focused on well-being, self-fulfillment and participation (Borzeix *et al.*, 2015), such as foresight or intrapreneurship, perceived as "buffer zones"[17] that allow a firm to secure the loyalty of their younger employees:

> As a former start-upper, I said "wait a minute, do you realize the impact you can make?" But in fact, BatiCorp E collaborators want to be able to dream of transforming the world, about the impact they're going to make, etc…[18]

Company directors count on their employees' dreams and desires, because the "principle of efficacy of [their] action […] resides in the ability to foresee and exploit trends to their own benefit." (Bourdieu & Boltanski, 1976, p. 54)

[14] Interview done 05/06/2019 with Nicolas F. (38), in charge of open innovation and collective innovation at BatiCorp E until end 2018, with a diploma from Science Po-Toulouse, specialized in project engineering and financing. He holds a masters in administration and communication from Toulouse University as well as a B.A. from the University of Montreal.

[15] Interview done 25/01/2019 with Wilfried C.

[16] Interview done 05/06/2019 with Nicolas F.

[17] Interview done 25/01/2019 with Wilfried C.

[18] Interview carried out on June 5, 2019 with Nicolas F.

The political socialization of directors of innovation

Structuring innovation has in fact become quite standard. At group level, a department heads a series of services disseminated among the different strata of the organization. At the top, responsible innovation combines with strategy. To be more precise, the top-level officials of the organization make durability and sustainability mandatory.

Though profiles vary, individuals' experiences as students and professionals are significantly labeled politically, particularly in economic diplomacy. Hugo T., in charge of BatiCorp's "prefiguration of innovation systems", director of "foresight", and since appointed director of the group's innovation program, is a good example. He began "by working on the interfaces between the public and the private spheres":

> I had huge projects of innovation, and connections with public policies on those subjects. I began work in telecommunications, in a sector for the Federation that represented telecom interests, which had huge stakes connected to the public sector.[19]

Hugo then rallied the economics department of the French Embassy in an Asian country, a service attached to the General Directorate of Public Finances, where he participated in accompanying French companies in their hunt for financial assistance and Government loans.

Hugo T.'s colleague, Wilfried C., was director of acceleration and entrepreneurship in the innovation department of BatiCorp. During a trip to the United States to deliver a paper in Berkeley, he met the French ambassador, who suggested he apply for a position as a "totally energy-and-environment-patented scientific attaché". In 2012, Wilfried was a diplomat in the United States.

Thomas' itinerary also reveals a strong interest in politics but, due to his marked activism, differently from his colleagues. His activity consisted at the time in "throwing out ideas, advocating them among government agents, various commissions and lobbying".[20] In point of fact, he was already involved in a university project with an ecological dimension, which shaped him professionally, halfway between expertise and political engagement:

> I was part of a program for the energy transition, Solutions Project, at the crossroads between the sciences, medias and politics. That's what developed into the "Green New Deal" of today, in America. At Stanford, I was in fact acting as Chief of staff. It was Mark Jacobson's idea, a professor at Stanford who was the first to develop plans for the State of Washington and the State of California, the United States, the world… U.S. states have worked separately on their own energy road

[19] Interview carried out on June 25, 2019 with Hugo T. (age 35), who trained as an engineer and directed the innovation program at BatiCorp. He is also a Science-Po Paris graduate in "Public Affairs".

[20] Interviews carried out on 28/10/2019 and 29/11/2019 with Thomas A., 33, in charge of open innovation and relations with the start-ups at BatiCorp. He was trained as an engineer and specialized in energy and nuclear engineering.

> maps, which then turned into the green new deal championed by Ocasio Cortez and people like that.[21]

A fourth member of the team went through AFNOR[22] (Cochoy, 2000), where she was in charge of developing a norm for "innovation management". She explains the political and economic reasons for developing such a norm:

> The European Commission figured they spend billions of money every year on innovation projects, but most of the time they fail, they're badly built. So what we need is criteria for the calls for proposals, to be able to evaluate the innovation projects of innovative companies better and allocate public funds better too.[23]

Functions of "interface" is the expression used by these individuals to describe their own interest in the public sector. Experiencing State institutions as students or professionals is the start of a socialization in the field of power and its mechanisms, particularly financial. Shuttling between the public and private sectors also allows them to acquire skills in the financial markets (Godechot, 2013) and in capital-investment funds (Benquet & Bourgeron, 2019). Their itineraries have therefore brought them into close contact with public innovation policies. Unsurprisingly, innovation managers judge, and speak their minds about, the policies applied, particularly in education, and readily discuss the training policies that will be applied to the future workforce when innovation is everything. That political socialization however does not fully explain why they adhere to the principles of responsible innovation. Their worldviews and values also allow us to understand their commitment to responsible innovation.

Responsibility, values and ethics among actors in charge of innovation

Interviews carried out with innovation managers are rife with anecdotes and statements extolling an industry that respects the environment by adopting measures of responsible and durable innovation. There are however variations due to differences in generation and fields of knowledge. The fact that innovation managers were trained as engineers is not without consequence. In their 2011 report, Christelle Didier and Kristoff Talin show that the profession of engineer is riddled with ethical differences (Didier & Talin, 2011). Most engineers (87%) consider their skills to be one of the conditions that keeps the planet running. Though there seems to be a tendency to under-estimate what is at stake ecologically, some innovation managers we spoke

[21] Interviews carried out on 28/10/2019 and 29/11/2019 with Thomas A.

[22] Association Française de Normalisation (AFNOR).

[23] Interview done on 09/07/2019 with Astrid K. (35), in charge of foresight in the innovation program of the BatiCorp company. She has a Masters in innovation management from a French university.

to – engineers among them – claim they have truly internalized the climate dimensions of human activity. This attitude may stem from a primary socialization extending into their higher education, particularly if they majored in "environment" or "energy":

> So in high-school I tried to discover what it is I wanted to do. I think environment attracted me more than energy… during the 2000s, my father used to say "we're going to have a problem – climate refugees… wars because of water, oil reserves are going down". You need energy for human activity but the idea is that the energy produced mustn't exhaust the resources. All that became clear in Engineering School.[24]

A professional sub-section of less-experienced engineers appears when observing recruitment in innovation departments or for work on environmental risks (Gadéa, 2015), echoing research by Goussard, Flocco and Petit, who note that some young engineers complain of "operational monotony" (2018). When an engineer opts for a position of innovation manager, it is a way of circumventing the more traditional production routines. Also, the political and ethical aspirations and commitments of these individuals may be at the root of their career choices. If that be the case, working for Bombardier or for a start-up boils down to a political choice:

> I was there for the specific project, i.e. I saw housing being produced on a small scale in Canada, which for me was a step in the right direction. Housing represents 50% of the primary energy consumed. There's a lot of talk about being vegetarian, it's very important, but only for 5 to 10%. In that case it meant working on 50% of the building, i.e. 50% of the equation of climate change. I looked up Tesla too, and other companies like that.[25]

Social and societal aspects also are important, even though they have been only partly defined (Bagattolli & Brandão, 2019). It is in that sense that innovation must be inclusive and take "human" factors into account. Innovation managers therefore become the guardians of the social consequences of innovations, by importing or producing criteria for responsible innovation. The value of empathy, inclusion and "co-innovation" feed into a normativity that they spread around the organization (Bedreddine, 2020b): "Co-innovation is the big theme… for everyone … with clients, with employees, for large and small firms, and start-ups."[26]

The various objectives that innovation managers mean to attain are in fact transversal. They claim a variety of fields of expertise that tend to overlap with other

[24] Interview done on 28/08/2018 with Maïlys C. (33), an engineer-researcher in the innovation department of the firm Énergéo. With a B.A. in physics and chemistry from Jussieu University, she specializes in environment and energy.

[25] Interviews done on 28/10/2019 and 29/11/2019 with Thomas A.

[26] Interviews done on 20/06/2018 with David L. (50), director of an innovation program on the intelligent city and director of open innovation on the internet in the Rés'O firm. He holds an engineering degree from a grande école. During his career he also obtained diplomas from major Business Schools in strategic management and in innovation management.

employees' "professional jurisdictions" (Abbott, 1988). They improvise, pretending to be in turn the champions of sustainable development and the professionals of personal development. What in fact is being challenged is technicist innovation. Innovation is made into a "transversal" and "holistic" fact (Chen *et al.*, 2018) that enjoins engineer-trained innovation managers to avoid the traditional representations where technique is topmost (Coutant, 2014). Technicism becomes the target, criticized for closing off a much vaster field of possibilities:

> Already it means transforming the world of labor, and through that, more globally, transforming the world. Working at BatiCorp E means playing in the field of cities, industry, construction and the well-being of the residents in those buildings… it means working everything that's going to be AI, real debates about ethics. At VivaTech two years ago, the concept of the Human as more than digital, is something we created with the woman who directed innovation, the woman who directed communications and the President. "What is the position, the posture that BatiCorp E defends as a firm and in what way is it different from the GAFA's..". I don't share a purely technological view of innovation, of the transformation of firms, of businesses and of the world…[27]

What we are seeing is the production of an ethos of innovation managers, that mixes economic considerations and the responsibility of an innovation. Thanks to the intercession of certain agents, the economic and political fields converge around worldviews and practices that in reality tend to justify the actions of innovation managers.

THE ENCOUNTER OF THE ECONOMIC AND POLITICAL FIELDS

Watching and foresight

The logics of financialization work differently in different social spaces (Darcillon, 2019; Faure *et al.*, 2019; Lebaron, 2015), even when responsible innovation calls for a projection that extends beyond the three years reference which is the benchmark for executive committees and stakeholders. Long-term calculations allow innovation managers to talk about the future by describing the positive and negative externalities of decision-making. At the same time, the short-termism of institutional investors (Plihon & Rigot, 2018) and the logics of large and small firms (Benquet *et al.*, 2019) draw further and further away from the need for far-sightedness that innovation managers demand. Yet their individual itineraries, values and a context favorable to long-term strategies, give innovation managers room to maneuver and negotiate their place. In that respect, they resemble other categories of individuals such as "finance prophets" (Pénet, 2019) or "promise builders" (Pollock & Williams, 2010); and in that sense, they organize the uncertainty.

[27] Interview done on 05/06/2019 with Nicolas F.

Innovation managers reinvest the results of their surveillance, from classical benchmarking to collecting legal data, through a complete re-examination of public action and of the systems that can benefit their group. These results trigger foresight, that consists in identifying one or several fields of possibilities. Since innovation managers claim to have 360-degree vision of the present and the future, they occupy a position of scout or truffle pig,[28] that ensures their situation as components in a firm's strategy. They emphasize the "trends" of the market, thereby attempting to orient the firm's production and activity. Innovation managers' ambition is to present innovation as a "social good", particularly by applying the concept of responsible innovation, which *in fine* permits hiding the negative effects described in their prospective work (Delvenne, 2017). Beyond collecting information for their collective activity, foresight consists in attracting the attention of the firm's upper echelons:

> We use it first of all to raise their awareness […] "hey, guys, we're going to lag behind if we don't act now". We were able to anticipate something and it was a real eye-opener for the sponsor, who at first said "what's climate resilience"[…] once we identified the 6 themes, there was a first exploratory phase, foresight, where the sponsor made up a group of the fifteen top managers of the different departments at BatiCorp around a theme, the aim being transversality, and the idea was to do it over 6-8 months – 4 days of work in the shop, to decide on a strategic positioning and a plan of action.[29]

Innovation managers then enter into an argument based on more or less scientifically corroborated data, but which serves their talk about the future by producing scenarios:

> A fantastic idea, I picked up another case about the CNES[30] who elaborated a prospective study on how to conquer space, and why *do* we want to go into space? Reality is going to up-end our hypotheses… for example, demographic growth, climate change, change in temperature, sure that's going to happen…[…] ask ourselves, "well, *why* go into space", so you imagine it and build those great scenarios. We bet on wars, we'll want to get away, so we'll want to go into space, politicians are going to say "it's hell on earth, we'll go into space and advance together…"[31]

Building scenarios for the future is therefore both operational and utilitarian. The long-term permits innovation managers to project their firm by fictionalizing (Petitprêtre *et al.*, 2019; Saint-Martin, 2019) and to rouse the employees. Simultaneously, that way of doing things fuels the activity of responsible innovation, without necessarily producing a final decision. Sessions of foresight with members of acting committees

[28] Interview done on 22/01/2018 with James R. (50), innovation coach and "catalyst" for the innovation department, part of R&D at Énergéo; graduated from a School of Engineering.

[29] Interview on 04/02/2019 with Stéphane Q. (50), who directs a department of innovation integrated in Énergéo's R&D. He went through the "operational", strategy and marketing. His present team is an accompaniment service for innovation. He holds an engineering degree from Centrale Supélec.

[30] Centre National des Études Spatiales.

[31] Interview on 09/07/2019 with Astrid K.

allow innovation managers to make sure their principles are given publicity, thus reaching the top levels of the firms.

In France, foresight as an activity was observed among the authorities as early as the late 1950s (Andersson & Prat, 2015), becoming more intensive during the early 1970s, with a view to "modernizing public action" (Jany-Catrice, 2019, p. 73). Foresight by the State consisted in anticipating and creating futuristic scenarios (Colonomos, 2014), to predict and build plans of action for what lay ahead. In fact, prospective methodologies promoted interactions between the State and the private sector thanks to strategies made possible by qualitative and quantitative tools (Andersson & Prat, 2015). Also, foresight appeared at the heart of firms well before innovation departments became a part of the system. Henceforth, taking a long-term view, as innovation managers of the large groups are wont to do, is part of the habitual functioning applied in the political and economic fields. Innovation managers apply a code common to the public as well as the private sectors, which deals with the future through foresight.

In firms, foresight is not new; the novelty resides rather in the fact that the function is taken in charge by a particular category of employees committed to innovation. It lies also in the nature of the problems that the foresight implemented by innovation managers proposes to deal with. Climate challenges, questions of inclusion at all levels or yet again mastering complex negative externalities, are taken up by innovation managers, whose specific itineraries speak in favor of taking extra-economic data into account. The arrival of Emmanuel Macron – an ally of business thanks to his own personal itinerary, his relationship to industry and his ideological stance (Offerlé, 2019) – as amplified the trend. Also, his election saw the advent of a staff up till then quite removed from politics and closer to the world of business (Michon, 2019). The context, favorable to industry and to private initiatives, comes with a new political awareness of the climatic and societal challenges at hand. Economics are not however pushed aside. Innovation managers juggle all at once with responsible innovation, communication strategies, and putting foresight in its economic context.

Meeting places

The encounter between public policies and the private sector occurs in various places. Political initiatives become part of a continuum (Pin, 2020) by generating the conditions of possibility for public-private interactions to emerge.

Commercialization services and the SATT network

During the years 2010, "commercialization services" were created in the universities in order to develop exchanges between the public sector and the "social-economic world". Clarysse A. directs such a service today in a French University. Previously, she

was in charge of technology transfer at the CNRS, principally through calls for projects:

> When I entered the CNRS, in 2005, at the time the ANR[32] was created, funding came from the Ministry and everyone did their research in their little corner, without necessarily caring about inno or transfer. Little by little, calls for projects showed an interest in the social-economic world […] not necessarily commercial, but fields like climate change, for example.[33]

Her present department is made up of five people gifted with a variety of skills, known as "commercialization specialists". Their mission is to economically enhance – or not[34] – (Lebaron, 2015, p. 4) the production of academic goods. The notion of commercialization springs straight out of unsuccessful political attempts (Flesia, 2006) at producing researcher-entrepreneurs, and once again questions university autonomy (Gibbons *et al.*, 1994). Commercialization concerns intellectual property, partnerships with the private sector and accompanying researcher-entrepreneurs.

At the same time, the Universities also set up full-fledged innovation departments, that supervise the work of commercialization specialists and do the job of communicating inside the University, to spread the principles of the innovation they wish to promote:

> The way I see my job is to facilitate the work of researchers who want to enter the business world, or create or commercialize the results of their research. If a researcher has an idea or an innovation and wants to go further, I help him or her develop the project and contact the right people.[35]

The departments accompany research-entrepreneurs, when they enter into partnerships in the private sector, on the legal and financial fronts, especially thanks to the contacts university commercialization and innovation departments entertain with big industry. The agents working in these departments point out the contradictions of their missions. They must both motivate and accompany "applied, or even very applied" research,[36] without however eliminating the responsible nature of the innovations:

> This morning in front of the Commission, I presented systems of the PACTE law that change the code of research, because of certain dispositions concerning

[32] Agence Nationale de la Recherche.

[33] Interview on 10/10/2019 with Clarysse A. (50). She directs the commercialization service of a French University, after having worked in the commercialization services public institutions. She also worked in a private company, in charge of barometric studies.

[34] These specialists insist on the fact that the products of academic research are not merely economic but can also be social or environmental; monetary profit is not their sole objective.

[35] Interview done on 10/05/2019 with Élise C. (35). She holds a diploma from a grande école where she did a 5-year curriculum in environmental sciences. She later obtained a Maters in innovation and commercialization engineering, developing her legal skills (patents, contracts, etc.).

[36] Interview on 10/10/2019 with Clarysse A.

> linkages between public and private, and the orientations, the national stakes involved in financing research, from competitivity to environmental issues or public health. Research concerns all of these. We need an economic model today, and to look for partners, because we need to answer calls for projects.[37]

The contradiction between climate issues and research through economic partnerships, leads commercialization specialists to take a relative view of the profitable nature of environmental research. According to Clarysse A., "if there was a market for the environment, we'd have known it". Economic profitability and responsible innovation thus seem barely compatible, despite the good intentions and political decisions in favor of an innovation that cares about the climate and social conditions. "Mercantilizing" science (Lamy & Shinn, 2006) remains aim number one for the public authorities, who since the 1960s have witnessed the economic opportunities offered by the circulation of techniques born of scientific research.

In 2010, an ANR call for projects on the transfer of technologies led to creating Technology Transfer Accelerator Offices (SATT[38]), private ventures whose mission it is to support public service researchers on the lookout to sell all or part of the results of their research. The innovation departments, the commercialization structures and, more recently, the SATT, participate in commercializing science and in introducing economic logics into French science; they participate in tightening the links between science and industry, which the agents of the political field have been wanting to see for decades.

Ecosystems, clusters and fairs

The environment created by the public authorities carries its lot of opportunities into the innovation departments of the large firms interested in externalizing part of their R&D. The public incubators and commercialization services described above are places innovation managers like to be, on the look-out for innovators to finance. The "Makerspace" (Anderson & Séac'h, 2012; Berrebi-Hoffmann *et al.*, 2018) and other "fablabs" (Bosqué, 2015) are favorite places for the "start-up scout" of a large firm. University or school incubators are valued because they are full of students trained in entrepreneurship and innovation methods (Chambard, 2013, 2020) that can benefit a student-entrepreneur and validate their year by a training period in entrepreneurship. The incubators in Universities and other public establishments also teem with potential "partners" for innovation managers.

[37] Idem.

[38] When they were created, the SATT (Société d'Accélération du Transfert de Technologies) received 1 billion Euros through the National Commercialization Fund (Fond national de valorisation), part of the ANR (Agence Nationale pour la Recherche), in order to "finance the commercialization of public research" (2010a, 2010b).

Open innovation, then innovation ecosystem (Bedreddine, 2020a, p. 75) are the concepts that actualize the interactions of innovation agents outside the firm. Innovation professionals patrol those places, on the lookout for the great idea, the dream team, the right product. Jacques, a start-up scout, explains what an ecosystem is:

> It's a place where you can spot and meet startups, they might be incubators, accelerators, investors, or maybe clusters, in France for instance there are a couple of clusters you can visit.[39]

The professional itineraries of a new political personnel are also a strong asset for innovation managers. The innovation ecosystem is an open book for minister Amélie de Montchalin:

> An ecosystem is a really informal gimmick, in a firm that was by definition very normative, hermetic, with limits… […] I can see innovation ecosystems around each group, where they think, well, that's our job, our mission is bound to evolve, so they surround us with people who're going to help us manage the changes.

Innovation managers are therefore allowed, or actually instructed, to circulate outside the organization, which sometimes exposes them to doing tasks considered unproductive. David L. uses the interview we did as an example to elucidate what drives open innovation, colored by serendipity and openness, where economic profitability is not necessarily the rule:

> I'll tell my boss I saw Samir Bedreddine, and he'll say that's great, that's fine, we talked. But you're not going to help me in my job, though the discussions we had and are going to have, we'll swap ideas, and maybe at some point you'll send me a note about the organization, well, innovation – we take it wherever we find it, and spread it around the organization.[40]

Fairs and events are spaces where people rub elbows, where people from all sorts of horizons congregate. VivaTech, the great innovation event in France, welcomes stands of all the large groups present on the French stock market. The French regions and universities are there too, to talk with the employees of the large groups, particularly with innovation managers. Élise, who belongs to the innovation department of one of the large French universities, also attends VivaTech:

> I meet industrialists during events at Spring[41], at VivaTech, aside from other appointments with them from time to time, but I also organize meetings directly

[39] Interview on 20/06/2018 with Jacques F., trained in a business school and start-up scout at Énergéo.

[40] Interview on 20/06/2018.

[41] Innovation Fair at Saclay.

> between researchers and industrialists, without first contacting any commercialization service.[42]

Awarding prizes to startups also implies that agents from various social spaces meet and exchange. Madame de Montchalin explains why those are opportunities for the private and public sectors to interrelate:

> What's important is that such occasions bring together very different people. Lab directors, financiers, public institutions, people like me, it means sharing things which normally have a future or could have one. You also get weak signals, fashions, fashions that can become trends…

In fact, each of these occasions are like places where the products of innovation of all kinds are recorded and consecrated, where *illusio* is "produced and reproduced" (Bourdieu, 2015a, p. 279-280). There, everyone speaks the same language, shares the same codes, without ever questioning what they are based on. Places of "hobnobbing", these "neutral places" (Bourdieu & Boltanski, 1976, p. 10) become the stage where the ideology of and belief in (Bourdieu, 1977) innovation – especially responsible innovation – are produced, asserted, disseminated and reasserted. We witness the construction of a shared universe, where material and symbolic goods are exchanged, whose supreme aim is the opening up and circulation of goods and individuals. Start-up juries, events of all sorts, partnerships or incubators, compose the constellation of French innovation. In the Triple Helix model, "hybrid organizations" can be found precisely in those shared and relatively autonomous spaces. Their main function is to promote innovation through the many objectives they defend. According to Champenois and Etzkowitz, these objectives are characterized by the fact their actions take place in many locations. As these authors say, these entities "integrate and combine elements from the various Triple Helix spheres in their institutional design, to promote innovation" (Champenois & Etzkowitz, 2018, p. 29), therefore participating in the activity of innovation agents by providing a field of professional possibles. From this point of view, innovation, and therefore responsible innovation, become the privileged objects for interactions between the academic and economic fields and the field of political power to take place.

Responsible innovation, a proxy for the interpenetration of the economic and political fields

Innovation becomes a shared code, a language that allows certain agents of the public and private sectors to interact. What surfaces, in reality, is a structural and

[42] Interview on 10/05/2019 with Élise C.

functional resemblance between the large companies' departments of commercialization and departments of innovation.

A structural and functional resemblance

The innovation departments of the large firms and the commercialization services in the universities and, more generally, in Academia, are the fruit of decisions made by the dominant actors in each field. Yet, a chronic sensation of being illegitimate plagues the agents in charge of innovation, as they move within their allotted fields, whose "nomos" represents a "supreme law", difficult to break without incurring punishment (Bourdieu, 2015b, p. 139). Individuals whose function it is to direct and accompany innovation find themselves in a quandary in fields governed by rules and mechanisms (which, incidentally, they call into question). That is why they incur mistrust on the part of other agents.

The paradoxical injunctions to which they are exposed give way to arrangements in which the heavy-handed laws of their field are a salient fact. Innovation agents do in fact sometimes make outlandish promises. Innovation managers therefore constantly remind everyone that the condition *sine qua non* of their action is profit-making and economic rationality. At the same time, commercialization specialists, as well as the other people involved in accompanying innovation in the public sector, corroborate and stress the need for objective and disinterested research, detached from any direct economic motive, not precluding, however, the idea of "going farther"[43].

Such paradoxical postures put the innovation departments of the economic and academic fields in a rather peculiar situation. Despite the support of their hierarchies, they are often pushed into the margins, due to behavior that appears eccentric with regard to the customary values and practices of their respective fields. The individuals who make up those services are therefore tempted to look elsewhere. They find themselves at the margins of their field, even sometimes straddling its borders. Trying to find justifications on the outside is not new and has been documented for other professions (Chiapello & Gitiaux, 2009). The originality here is the way the positions occupied by agents of the different fields adjust. Individuals come to the fore whose internal and external legitimacy mix and blossom in their shared spaces. They play with the rules and limits of their fields and operate in the nooks and crannies. Representatives of private and public innovation might be called *agents of the interstices*, that Etzkowitz and Champenois call "boundary spanners", or to coin a neologism, *interstitial agents.*

[43] Interview on 10/10/2019 with Clarysse A.

As a result, innovation agents in each field share an ensemble of practices and world views. Their shared values of openness, freedom and the abolition of borders – especially scientific borders (Gibbons *et al.*, 1994) – are their common, normative platform. They use the same language, which forms and transforms discussions in fields whose basic principles are at first sight quite far apart. This is what imparts its significance to the meeting places mentioned above.

Entry gates

The production of students (Chambard, 2020), first, the production of research (Lanciano-Morandat, 2019), second, and finally, the creation of physical and normative areas, are what permit the Public Authority to produce the offer of innovation in France. Despite the ongoing climate of austerity (Guilbert *et al.*, 2019), the systems for accompanying individual and collective initiatives of entrepreneurs and firms are plethoric. The BPI (Public Investment Bank), research tax credit and all sorts of State subventions, constitute areas of interpenetration that make the State a major actor where investing in innovation in France is concerned. The legal arrangements defining corporate social responsibility point in the same direction and illustrate the power of "the brain-washing done by the State" furthered by the public authorities (Bourdieu, 2014, p. 123), who thereby reassert their role as producers of reality (Bourdieu, 2005), through a "double social construction"[44] of the market (Bourdieu, 1997, p. 49). Responsible innovation serves that end.

That is how innovation and its supporters in the economic, political and academic fields work at building *entry gates*, by *consecrating* symbolic and material goods, in an exchange ranging from the right vocabulary to correct behavior, passing through objects of inter-comprehension. By occupying the position in charge of organizing the firm's public relations, innovation agents become the gate keepers of their field. On their common meeting grounds, responsible innovation is the main theme. Individuals act in accordance with the rules that govern their field, and transcend them by adopting foreign logics. For innovation agents in the public sector, personal stakes and the common good are not mutually exclusive. Producing knowledge and skills aimed at a better understanding of the economic, social, biological or physical worlds, does not prohibit economicist side-stepping. On the side of innovation managers, bypassing economic rationality balances out their insistence on the economic virtues of what they say and what they do. On both sides, aims are now hybrid.

Innovation managers' job then consists in making their firms' employees "sensitive" to the issues of responsibility and sustainability. Responsible innovation,

[44] A double construction, in the sense that the State produces individual aspirations and a field of possibility (laws, financing, subventions), in which those "systems of individual preferences" can prosper (Bourdieu, 1997, p. 49).

which falls within the scope of both economics and politics, consequently appears as the central theme of discussion between the public authorities and economists. The responsibility of innovations becomes a proxy for their interactions, and the star product, the main object shared in the political and economic fields.

Responsible innovation has thus become one of the main modes of interaction between the two fields. The phenomenon is not new and has been described as the "Triple Helix" (Etzkowitz & Zhou, 2017), a configuration where the University, industry and public authorities join up. However, the transformation operated in France permits us to introduce new elements with relation to a concept sometimes described as "ready-made ideas about science" (Shinn, 2002). For, in reality, creativity in places dedicated to innovation and related tasks, in both the public and private sectors, is the result of planning by the leaders in the fields and also of the desire voiced by certain agents. That is how individual aspirations join up with the obligation to conform imposed from the top in the fields.

CONCLUSION

Economics and politics are transformed by the action of individuals possessed of specific personalities and values. We witness the difficult construction of a professional ethos that places freedom and openness at the center. Discourses contain a hegemonic will that sees everything through the lens of innovation. Innovation agents working in the interstices of organizations and fields are at odds with the rules of expertise. Gifted with multiple resources, but not enough to carry weight in the fields dominated by certain types of capital, they have trouble asserting themselves within the firms.

At the same time, innovation increasingly appears as an activity born of private initiative. Nevertheless, responsible innovation is an important stake in regulating both the public and the private sectors. Innovation agents in the economic and political fields try to make profitability and responsibility work together. A contradictory mantra, yet responsible innovation emerges as a facilitator in the effort to reconcile the two, observed both in legal and professional practices. The "firms' *raison d'être*" intersects with innovation managers' values, participating in the mechanics of interaction between fields. The public authorities pick up the term innovation and work on its semantics. The word refers to the transformation of both the firms and public research. Furthermore, we see the birth of an "entrepreneurial man" and a "new conception of the market" (Dardot & Laval, 2010b), at a time when neoliberalism was emerging as the "new reason of the world" (Dardot & Laval, 2010a). The making of new individuals, how they relate to the self, the group, the State and the market, seems to

be one of the purposes of all innovation and of the precepts of those who defend it. Demands for less State intervention mingle with the massive presence of public authorities in French innovation.

This article in fact raises the issue of the fields' loss of autonomy, indirectly due to innovation agents. The autonomy of fields depends here on the autonomy of professionals (Sapiro, 2019), which we have attempted to describe. Are we witnessing the emergence of a field of innovation claiming its own market and its own mechanisms? As things stand, we observe an interdependency that hardly supports such a development. Distancing from the all-economic and the all-public – hybridity *par excellence* promoted by responsible innovation – remains at the margins of the fields and their modes of functioning, which though established, nevertheless might change. The "startup nation", promised by presidential candidate Emmanuel Macron, who set as a condition the possibility of replicating the California model (Etzkowitz, 2019), seems weakened by inconclusive results. The myth is falling by the wayside, as today, the phrase is mainly employed pejoratively, and even its representatives express doubts as to its validity.[45]

REFERENCES

Abbott, A. (1988). *The System of Professions – An Essay on the Division of Expert Labor*. University of Chicago Press.

Amable, B., Chatelain, J.-B., & Ralf, K. (2006). Nantissement des brevets et croissance des innovations. *Revue d'economie politique*, 116(4), 523-540.

Anderson, C., & Séac'h, M. L. (2012). *Makers : La nouvelle révolution industrielle*. Pearson.

Andersson, J., & Prat, P. (2015). Gouverner le « long terme ». *Gouvernement et action publique*, 4(3), 9-29.

Bagattolli, C., & Brandão, T. (2019). Counterhegemonic Narratives of Innovation. *NOvation – Critical Studies of Innovation*, 1(June), 67-105. http://www.novation.inrs.ca/index.php/novation/article/view/6

Bargel, L. (2014). Apprendre un métier qui ne s'apprend pas. Carrières dans les organisations de jeunesse des partis, Learning a craft that can't be learned. Careers within party youth organizations. *Sociologie*, 5(2), 171-187. https://doi.org/10.3917/socio.052.0171

Bedreddine, S. (2020a). Des grandes entreprises et des start-up : Logiques d'interactions, pratiques de contrôle. *Savoir/Agir*, 51(1), 69-77.

Bedreddine, S. (2020b). Le Design Thinking : Dispositif de gestion de la créativité dans les organisations et outil de hiérarchisation des salariés. *Cahiers internationaux de sociologie de la gestion*, 22, 11-36.

[45] On December 6, 2019, Olivia Grégoire, an elected representative of the French Presidential majority (also appointed minister) rejected the term startup nation employed ironically by the host of a TV program on a French news channel.

Bedreddine, S., & Noûs, C. (2021). Dire et diriger l'innovation pour transformer l'organisation du travail. Mots. *Les langages du politique*, 126(2), 15-32.

Bennett, S., Maton, K., & Kervin, L. (2008). The 'Digital Natives' Debate : A Critical Review of the Evidence. British Journal of Educational Technology, 39(5), 775-786. https://doi.org/10.1111/j.1467-8535.2007.00793.x

Benquet, M., & Bourgeron, T. (2019). Accumuler le capital. *Actes de la recherche en sciences sociales*, 229(4), 46-71.

Benquet, M., Bourgeron, T., & Reynaud, B. (2019). Économie politique de la financiarisation. *Actes de la recherche en sciences sociales*, 229(4), 4-13.

Bereni, L., & Prud'Homme, D. (2019). Servir l'entreprise ou la changer ? *Revue francaise de sociologie*, 60(2), 175-200.

Berrebi-Hoffmann, I., Bureau, M.-C., & Lallement, M. (2018). *Makers*. Le Seuil.

Borzeix, A., Charles, J., & Zimmermann, B. (2015). Réinventer le travail par la participation. Actualité nouvelle d'un vieux débat. Introduction. *Sociologie du travail*, 57(1), 1-19. https://doi.org/10.4000/sdt.1770

Bosqué, C. (2015). Enquête au cœur des FabLabs, hackerspaces, makerspaces. Le dessin comme outil d'observation. Techniques & Culture. *Revue semestrielle d'anthropologie des techniques*, 64, 168-185. https://doi.org/10.4000/tc.7579

Bourdieu, P. (1977). La production de la croyance [contribution à une économie des biens symboliques]. *Actes de la recherche en sciences sociales*, 13(1), 3-43. https://doi.org/10.3406/arss.1977.3493

Bourdieu, P. (1997). Le champ économique. *Actes de la Recherche en Sciences Sociales*, 119(1), 48-66. https://doi.org/10.3406/arss.1997.3229

Bourdieu, P. (2005). *The social structures of the economy*. Polity.

Bourdieu, P. (2012). *Sur l'État. Cours au Collège de France*. Le Seuil.

Bourdieu, P. (2014). *Raisons pratiques (réédition). Sur la théorie de l'action*. Points.

Bourdieu, P. (2015a). *Les règles de l'art. Genèse et structure du champ littéraire*. Points.

Bourdieu, P. (2015b). *Méditations pascaliennes (édition revue et corrigée)*. Points.

Bourdieu, P., & Boltanski, L. (1976). La production de l'idéologie dominante. *Actes de la Recherche en Sciences Sociales*, 2(2), 3-73. https://doi.org/10.3406/arss.1976.3443

Chambard, O. (2013). La promotion de l'entrepreneuriat dans l'enseignement supérieur. Les enjeux d'une création lexicale. *Mots. Les langages du politique*, 102, 103-119. https://doi.org/10.4000/mots.21374

Chambard, O. (2020). *Business Model*. La Découverte.

Champenois, C., & Etzkowitz, H. (2018). From boundary line to boundary space : The creation of hybrid organizations as a Triple Helix micro-foundation. *Technovation*, 76-77, 28-39.

Chen, J., Yin, X., & Mei, L. (2018). Holistic Innovation : An Emerging Innovation Paradigm. *International Journal of Innovation Studies*, 2(1), 1-13. https://doi.org/10.1016/j.ijis.2018.02.001

Chiapello, È., & Gitiaux, F. (2009). Les responsables développement durable des grandes entreprises. Parcours, engagement et représentations. *Revue de l'organisation responsable*, 4(1), 43-53.

Cihuelo, J. (2020). Le temps de travail des cadres à l'épreuve d'un dispositif de gestion de la créativité. Temporalités. *Revue de sciences sociales et humaines*, 31-32. https://doi.org/10.4000/temporalites.7615

Cochoy, F. (2000). De l'« AFNOR » à « NF », ou la progressive marchandisation de la normalisation industrielle. *Réseaux. Communication - Technologie - Société,* 18(102), 63-89. https://doi.org/10.3406/reso.2000.2258

Colonomos, A. (2014). *La politique des oracles : Raconter le futur aujourd'hui*. Albin Michel.

Coutant, H. (2014). La « technique » comme activité ou comme représentation partagée. *Annales des Mines - Gerer et comprendre*, 117(3), 49-58.

Darcillon, T. (2019). Finance et inégalités. *Actes de la recherche en sciences sociales*, 229(4), 72-85.

Dardot, P., & Laval, C. (2010a). *La nouvelle raison du monde*. La Découverte.

Dardot, P., & Laval, C. (2010b). L'homme entrepreneurial. In P. Dardot & C. Laval, *La nouvelle raison du monde* (pp. 219-241). La Découverte.

Delvenne, P. (2017). Responsible research and innovation as a travesty of technology assessment? *Journal of Responsible Innovation*, 4(2), 278-288. https://doi.org/10.1080/23299460.2017.1328653

Didier, C., & Talin, K. (2011). *Les ingénieurs et l'éthique* (p. 30) [Report]. Association des Ingénieurs et scientifiques de France. https://halshs.archives-ouvertes.fr/halshs-00784769

Dolez, B., Fretel, J., & Lefebvre, R. (2019). *L'entreprise Macron*. Presses Universitaires de Grenoble.

Edgerton, D. (2013). *Quoi de neuf ? Du rôle des techniques dans l'histoire globale*. Le Seuil.

Edwards, P. N., & Hecht, G. (2005). Les techniques de la guerre froide dans une perspective mondiale : Le nucléaire et l'informatique comme systèmes technopolitiques. In D. Pestre (Éd.), *Deux siècles d'histoire de l'armement en France : De Gribeauval à la force de frappe* (p. 167-178). CNRS Éditions.

Etzkowitz, H. (2002). *MIT and the Rise of Entrepreneurial Science*. Routledge.

Etzkowitz, H. (2019). Is Silicon Valley a global model or unique anomaly? I*ndustry and Higher Education*, 33(2), 83-95. https://doi.org/10.1177/0950422218817734

Etzkowitz, H., & Zhou, C. (2017). *The Triple Helix : University–Industry–Government Innovation and Entrepreneurship* (2nd edition). Routledge.

Faure, S. B. H., Joltreau, T., & Smith, A. (2019). Qui gouverne les grandes entreprises de la défense ? Contribution sociologique à l'étude des capitalismes en France et au Royaume-Uni. *Revue internationale de politique comparee,* 26(1), 11-45.

Flesia, E. (2006). Valorisation de la recherche, innovation et création d'entreprises. *Geographie, economie, societe,* 8(1), 149-158.

Freeman, C. (1995). The 'National System of Innovation' in historical perspective. *Cambridge Journal of Economics*, 19(1), 5-24. https://doi.org/10.1093/oxfordjournals.cje.a035309

Gadéa, C. (2015). Logiques professionnelles et problématiques environnementales. Introduction au Dossier « Enjeux environnementaux et dynamique des groupes professionnels ». *SociologieS*, 1-8. https://doi-org.inshs.bib.cnrs.fr/10.4000/sociologies.5113

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., & Scott, P. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies.* SAGE Publications Ltd.

Gingras, Y., & Villedieu, Y. (2010). *Propos sur les sciences*. Liber.

Godechot, O. (2013). 17. Concurrence et coopération sur les marchés financiers. Les apports des études sociales de la finance. In F. Vatin & P. Steiner, *Traité de sociologie économique* (p. 635-670). Presses Universitaires de France.

Goussard, L., Flocco, G., & Petit, S. (2018). Les ingénieurs face aux transformations du système productif. Des réactions contrastées aux légitimations partagées. In A. Dérouet & S. Paye, *Les Ingénieurs, unité, expansion, fragmentation (XIXe et XXe siècles)* : Tome I, *La production d'un groupe social* (p. 275-293). Editions Classiques Garnier.

Granovetter, M. (2017). 2. The Impact of Mental Constructs on Economic Action : Norms, Values, and Moral Economy. In Society and Economy : Framework and Principles (p. 26-55). The Belknap Press.

Guilbert, T., Lebaron, F., & Peñafiel, R. (2019). Introduction. Discours austéritaires et discours néolibéral. *Langage et societe*, 166(1), 9-29.

Jany-Catrice, F. (2019). Transformations de long terme dans l'évaluation des politiques publiques. D'une planification politique à une légitimation scientifique. *Actuel Marx*, n° 65(1), 67-80.

Joly, M. (2017). *L'Europe de Jean Monnet.* Cnrs.

Jonas, H. (2013). *Le principe de responsabilité : Une éthique pour la civilisation technologique*. Flammarion.

Lamy, E. (2020). *Les politiques françaises de « startupisation » de la science*. Savoir/Agir, 51(1), 23-32.

Lamy, E., & Shinn, T. (2006). L'autonomie scientifique face à la mercantilisation. *Actes de la recherche en sciences sociales*, 164(4), 23-50. https://doi.org/10.3917/arss.164.0023

Lanciano-Morandat, C. (2019). *Le travail de recherche. Production de savoirs et pratiques scientifiques et techniques*. CNRS.

Lebaron, F. (2015). Injonction comptable et révolution culturelle à l'Université. *La nouvelle revue du travail*, 6. https://doi.org/10.4000/nrt.2177

Lebaron, F. (2016). 2. La croyance économique dans le champ politique français. *Regards croises sur l'economie*, 18(1), 32-44.

Maitre, J., & Bourdieu, P. (1994). *L'autobiographie d'un paranoïaque*. Economica.

Michon, S. (2019). L'entreprise chevillée au corps. L'entrée d'une « société civile » entrepreneuriale à l'Assemblée. In B. Dolez, J. Fretel, & R. Lefebvre, *L'entreprise Macron* (p. 217-228). Presses Universitaires de Grenoble.

Née, É., Oger, C., & Sitri, F. (2017). Le rapport : Opérativité d'un genre hétérogène. *Mots. Les langages du politique*, 114, 9-24.

Negroponte, N. (1996). *Being digital* (1. Vintage Books ed). Vintage Books.

Offerlé, M. (2019). « Les patrons » ou « des patrons » avec Emmanuel Macron. Capitaux entrepreneuriaux et capital politique. In B. Dolez, J. Frétel, & R. Lefebvre, *L'entreprise Macron* (p. 79-92). Presses Universitaires de Grenoble.

Paredes-Frigolett, H. (2016). Modeling the effect of responsible research and innovation in quadruple helix innovation systems. *Technological Forecasting and Social Change*, 110, 126-133. https://doi.org/10.1016/j.techfore.2015.11.001

Pénet, P. (2019). Les prophètes de la finance. Contester et refaire l'époque. Tracés. *Revue de Sciences humaines*, 36, 59-79.

Petitprêtre, B., Perseil, S., & Pesqueux, Y. (2019). *La réalité de la fiction : Ou des relations entre fiction, narration, discours et récit*. Editions L'Harmattan.

Pin, C. (2020). La gouvernance territoriale de l'innovation. *Gouvernement et action publique*, 9(1), 57-85.

Plihon, D., & Rigot, S. (2018). Pourquoi manque-t-on d'investisseurs à long terme ? *Revue d'economie financiere*, 130(2), 113-128.

Pollock, N., & Williams, R. (2010). The business of expectations : How promissory organizations shape technology and innovation. Social Studies of Science, 40(4), 525-548. https://doi.org/10.1177/0306312710362275

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). Free Press.

Saint-Martin, A. (2019). Science-fiction et futurologie de la colonisation martienne. Espaces des possibles, régimes de croyances et entrecroisements. *Socio. La nouvelle revue des sciences sociales,* 13, 45-69. https://doi.org/10.4000/socio.7681

Sapiro, G. (2019). Rethinking the Concept of Autonomy for the Sociology of Symbolic Goods. *Biens Symboliques*, 4, 1-50.

Shapin, S. (2006). *The Man of Science*. Cambridge University Press.

Shinn, T. (2002). Nouvelle Production du Savoir et Triple Hélice : Tendances du prêt-à-penser les sciences. *Actes de la recherche en sciences sociales*, 141(1), 21-30. https://doi.org/10.3406/arss.2002.2815

Thébaud-Mony, A. (2014). *La science asservie*. La Découverte.

Ughetto, P. (2018). *Organiser l'autonomie au travail : Travail collaboratif, entreprise libérée, mode agile... L'activité à l'ère de l'auto-organisation.* FYP éditions.

Broadening our horizons. Digital technology, metatechnologies, and their implications for responsible innovation

Vincent Bryce[1]*, Tonii Leach, Bernd Stahl*** and Laurence Brooks******

* *University of Nottingham*

** *de Montfort University*

*** *de Montfort University*

**** *de Montfort University*

ABSTRACT

This paper argues that responsible innovation discourses must consider the changing nature of digital innovation, if they are to stand a chance of steering the development of technology towards democratically-acceptable ends. It explores the extent to which foundational narratives of Responsible (Research and) Innovation (RRI) consider problematic features of metatechnologies – defined here as "core information technologies upon which others are based, and whose use vastly expands the degrees of freedom with which humans can act in the social and material worlds" – and implications for responsible innovation discourse in the digital age. The study finds that references underpinning paradigmatic RRI accounts include digital and metatechnology examples, albeit briefly in some cases, somewhat reinforcing the validity of seminal RRI accounts in the context of new and emerging digital technologies with metatechnological attributes. The need for additional reflection on the problematic implications of digital technologies for RRI is identified, for example with respect to distributed development, and recombinant and network-level effects. The paper concludes that the continuing value of RRI as a discourse to society will depend on researchers' and practitioners' awareness of the potential of these technologies for cascading, downstream innovation.

Keywords: Responsible Innovation; Responsible Research and Innovation (RRI); Digital Technology; Metatechnology; Critical Hermeneutics.

Proposal Submitted 27 August 2020, Article Received 9 April 2021, Reviews Delivered 21 May 2021, Revised 12 June 2021, Accepted 13 October 2021, Available online 28 February 2022.

1 Acknowledgement: The corresponding author is supported by the Horizon Centre for Doctoral Training at the University of Nottingham (UKRI Grant No. EP/S023305/1) and by ORBIT, the Observatory for Responsible Research and Innovation in ICT (orbit-rri.org).

INTRODUCTION

Responsible Research and Innovation[2] (RRI) provides a framework to align innovation with societal need, and rests on a lattice of assumptions regarding the nature of innovation, the ability to anticipate effects, and the extent to which potentially impacted stakeholders can be meaningfully engaged.

Different RRI accounts frame the need for processes to align innovation with societal needs in different ways. Within an overall understanding of RRI as an interpretively flexible umbrella term (Rip & Voß, 2013), Timmermans and Blok (2018) identify four foundational perspectives originating from the work of von Schomberg (e.g. 2013), Stilgoe *et al.* (2013), van den Hoven (2013) and that of the European Commission (2012a). Each account is developed with reference to different types of technology.

While the challenges of epistemological insufficiency, conflicting stakeholder interests, and limits to transparency identified by Blok and Lemmens (2015) are frequently cited as general problems in applying responsible innovation theory to practice, each of these four accounts is developed with reference to specific technology examples. Because some types of technology pose additional challenges in relation to responsible innovation (for example digital technologies; Stahl, 2015; Jirotka *et al.,* 2017), it follows that if foundational perspectives of RRI did not consider these types, the frameworks they set out may not have fully considered their associated problems.

This may be particularly the case for emerging technologies (those which have been invented, but whose details and potential uses, researchers and end users are not yet fully aware of; Kendall, 1997), and in particular those which enable a very large number of potential uses through onwards innovation. Cressman (2020, p. 21) neatly describes the contextual significance of this in defining innovation as "a background of assumptions and attitudes through which technology is thematized and made meaningful, providing a context that directs technological society towards particular ends while simultaneously foregoing other ends" – in other words, as the assumptions, attitudes, and meanings associated with emerging technologies change, the scope of potential uses and context for directing these technologies to particular ends also changes.

To explore the extent to which foundational RRI accounts are anchored in underlying assumptions about contemporary technologies, this paper addresses the

2 While a distinction can be made between a policy-based concept of Responsible Research and Innovation and a broader Responsible Innovation discourse, as the terms emerged in parallel and have common features (Owen & Pansera, 2019) they will be used interchangeably in this paper.

question 'to what extent do the examples cited by foundational narratives of RRI consider problematic features of digital and metatechnologies, and what are the implications for the foundational assumptions of responsible innovation discourse in the context of societies' increasing use of digital technologies?'.

We begin by assessing the need to consider technologies as foreground phenomena. The paper then develops a concept of 'metatechnology' to consider ways in which some technologies may have a qualitatively greater potential to impact society, and then uses the case of digital technologies to explore the ways in which these features are problematic from a responsible innovation perspective. These aspects are then explored through an analysis of foundational accounts of responsible innovation using a critical hermeneutic approach, with specific attention to the conceptions of technology they reference.

THE NEED TO BRING TECHNOLOGY TO THE FOREGROUND

The tendency of philosophy of technology narratives to consider technology *en bloc* has led to calls to make particular technologies 'foreground phenomena', to reflectively analyse them in such a way as to illuminate features of the broader phenomenon of technology itself. Von Schomberg and Blok (2019, p. 7-8; p. 13) highlight the need to consider particular technological innovations to understand their effect in shaping moral decisions, and to enable us to evaluate the sense in which some innovations differ from each other and are either more or less ethically acceptable, societally desirable, and inherently controllable than others.

The printing press provides an historic illustration. While scholars saw block printing's potential to increase the circulation of religious works, and even expected it "would strengthen religion and enhance the power of monarchs" (Meyrowitz, 1995, p. 41), the potential for social reform and distribution of 'innovative' (in the sense of 'subversive') pamphlets from unregulated presses was not appreciated until the technology was widely available. For the purposes of our argument, the ability of this technological artefact to enable further innovation in the types of material that could be produced, the ways they could be distributed, and the social and other innovations resulting from the distributed material is a feature that differentiates it from others.

In defining 'technology', Arthur (2009, p. 18) helpfully distinguishes potentially different meanings, as:

1. *A means to fulfil a human purpose*
2. *An assemblage of practices and components*

3. *The entire collection of devices and engineering practices available to a culture*

From this perspective, any effort to consider the responsibility implications of a technology must consider attributes it possesses that may influence its use in practice – innovation produces technologies, but technologies can also enable innovation. This position knowledges both that technologies have a tangible form or substance (whether as artefact, or practice) which affects its potential for use, and that this form may be adapted by users in different contexts.

Within the context of emerging technologies, to an extent the challenge of RRI assessment is one of clarity over the nature of what we are trying to bring to the foreground. The concept of 'artificial intelligence' is a case in point – purported innovative technologies may be too vague to be the subject of analysis, and may be better understood as category labels, even as category error (the 'AI effect' – Hainlein & Kaplan, 2019). Blok (2020, p. 17-18) references examples of digital technology in highlighting the need to consider the potential of new technologies to create new world orders, beyond the creation of new artefects and services. This indicates the difficulty of assessing implications of innovative processes before specific uses have developed, and the need to consider emerging, and often constantly changing combinations of new practices to assess the potential for new effects.

TOWARDS A CONCEPT OF METATECHNOLOGY

Because assessing the implications of technologies involves consider their intrinsic features as well as the uses they are put to, features that affect the extent to which they can be reconstituted in use are particularly relevant (Orlikowski, 1992, p. 15). A hermeneutic perspective invites us to consider the potential for different meanings to be invested in an object, to more fully understand how it might impact in practice.

Blok (2020) highlights the idea that, unlike pre-existing understandings of innovation which from either an economic or philosophical perspective are based on identifiable commercial applications or methods of production, disruptive innovations are instead associated with the unknown, and draws our attention to the creation and evolutionary stages of technologies prior to market adoption.

For this reason, we need to consider how the properties of some technologies may predispose them towards different imagined, and potentially as-yet-unimagined uses. Feenberg (2017, p. 137) articulates this in the thought that technology is "not only artifactual, but also refers to the question of what we do when we envisage the world with a technical intention".

Theories of 'disruptive innovation', while contested, liable to reification, and potentially also rationalisation of a fear-driven aspect of commercial imitation, provide a starting point for us to articulate a concept of technology that differentiates those with more limited, and more extensive potential to be reimagined and impact in different ways.

Abernathy and Clark introduce the concept of 'transilience' (1985), defining this as "the capacity of an innovation to influence the established systems of production and marketing" (Abernathy & Clark, p. 3). With an explicitly commercial focus on the US auto industry, they distinguish types of innovation based on the extent to which markets, rather than just producer competences are disrupted. In their analysis, two subtypes are of interest – 'revolutionary' innovations such as radically more powerful car engines disrupt existing competence without creating new customer-market linkages, and 'architectural' innovations which result in changes to established systems of production, the creation of new industries and the reorganisation of old ones.

Utterback (1994, 1996) similarly defines 'radical innovations' as those which can 'sweep away' skills, knowledge, production techniques, and industrial equipment. This connotes a change in outlook, later explored by Bessant (2013) whose concept of 'paradigm innovation' is based on the extent to which mental models of production are changed. Christensen (1997, 2015) identifies two preconditions for 'disruptive technology innovations' – significant changes to attributes of existing products, and significant incentives for new business models compared to the old.

Brynjolffson and McAfee (2014) demonstrate the ways digital technologies and in particular, their evanescent marginal cost of reproduction create these preconditions. For Kodak, digital technologies created a double disruption – digital flash memory provided a more cost-effective replacement to film camera, but within a short period of time substantially replaced the practice of printing copies of pictures with the ability to share memories through social media.

Beyond market-oriented conceptions of radical innovation, we can see from a historic perspective that a number of technologies created the conditions for significant impact through adaptation to further uses and cascading innovation, and from this perspective could be assessed as 'radical', from fire, the compass, and gunpowder, to the printing press and steam power. We can see in each case that impact follows not so much the development of a method or artefact, but its association with expanded uses – the observation that China discovered gunpowder and the compass but applied them to fireworks and interior design is relevant here.

It follows that technologies will have more potential to impact if they have properties that increase the likelihood or extent to which they can be adapted to

different contexts and uses – in the hermeneutic sense, in their potential to take on new meanings.

The concept of 'metatechnology' provides a linguistic vehicle to distinguish innovations on this basis. While mutatory aspects were explored in earlier discussions of the philosophy of technology – for example Jonas (1979, p. 38) discusses "the Promethean enterprise of modern technology" – and in nanotechnology debates that informed RRI discourse, the first apparent use of the term is by Bross (1981). The sense of 'meta' here is of oversight and safety, through use of systems to prevent industrial accidents and enhance societal benefits of mammography – in effect, technologies to govern other technologies.

Vallenilla (1999) proposed the term to denote the purpose attached to the development or application of a technology – in Aristotelian terms, its 'final cause', for innovations that...

> ...seek to overcome the traditional anthropomorphic, anthropocentric, and geocentric limits of all previous technology… that often operate outside the bounds of human or natural powers and forms of sensation (e.g., nuclear energy and radar), go beyond enhancing human life as it is given (as with many unintended consequences of technology such as global climate change), or affect not just the earth but even the moon and planets. (Vallenilla, 1999, p. 411

This transhumanistic conception of metatechnology is of limited use, as we can attach an intention to a technology that may exceed its capabilities – I might intend to travel to the moon in a steam-powered rocket, but I am unlikely to reach the outer atmosphere. Similarly, we may not have this intention for a technology, but it may have far-reaching implications, as in the case of the ARPANET.

Braman (2004) defines metatechnologies in relation to their processing potential, and their potential range of outputs:

> Meta-technologies involve many processing steps, and there is great flexibility in the number of steps and the sequence in which they are undertaken. They can process an ever-expanding range of types of inputs and can produce an essentially infinite range of outputs… Their use vastly expands the degrees of freedom with which humans can act in the social and material worlds, and characterizes the postmodern world. (Braman, 2004, p. 5)

This account sees metatechnologies as always informational in nature. The concept is assessed in a historical context as convergences of communication with other materials and social processes, in the first case through the emergence of writing. Braman sees the modern information society and its harmonised information and communication systems as creating a situation in which "information flows have structural effects as powerful as those traditionally associated with the law" (Braman, 2004, p. 35-36), with the consequence that the ability to shape these flows – whether through their design, commercialisation or control – confers significant power. This

definition resonates with contemporary discussions of the political power of social media companies.

Mitcham (1995, p. 16), while citing Vallenilla, similarly highlights increasing interconnectivity and ubiquity in postulating a concept of metatechnology that "steps beyond the specific autonomies of modernity", although his subject is technology writ large and metatechnology is indicated as a replacement for the concept of technology that has gone before rather than a subcategory.

In elaborating our idea of metatechnology we can draw on the earlier concept from economics of 'General Purpose Technologies' (GPT), "deep new ideas or techniques that have the potential for important impacts on many sectors of the economy" (Wright, 2000, p. 161). This economics-focussed conception is elaborated by Jovanovic and Rousseau (2005) who identify 'pervasiveness', 'improvement' (in the sense of continuing and cascading improvements, such as those which reduce use costs), and 'innovation-spawning' as the characteristics of a GPT. While 'pervasiveness' may be better understood as an emergent quality, and the authors suggest that beyond these attributes GPTs do not necessarily differ from other technologies, these features, and the extremely broad examples cited as GPTs of 'electricity' and 'information technology' introduce a sense in which we are identifying as significant those technologies which enable the creation of others.

This progenitive aspect is picked up by Glazer (2007, p. 120) who defines metatechnologies as "the core technologies on which innovations are based", albeit identified in relation to marketable product characteristics, and by Romer (2009), who uses the phrase 'meta-ideas' to describe those which support the production and transmission of other ideas.

A different method of assessing what we might call the emancipatory potential of innovations is discussed by Edwards-Schachter (2018), whose concept of disruptive innovation, in contrast to Christensen's (1997) sees disruptive potential as a property of the person or organisation innovating as well as of the item being innovated. A technology not disruptive in one context, may be in another. The emancipatory or enabling aspect of a technology – which we could see as the ease with which it can be applied by new users, and which economists might see in terms of low barriers to entry – is also discussed in the concept of 'enabling technologies' that underpin 'Industrie 4.0' (Kagermann, 2011; Culot *et al.*, 2020). While this concept has been adopted as part of EU industrial strategy (European Commission, 2018), the concept of Key Enabling Technologies (KETs) used here is defined instrumentally, with reference to policies aimed at improving regional competitiveness:

> [KETs] enable innovation in process, goods and service innovation throughout the economy and are of systemic relevance. They are multidisciplinary, cutting across many technology areas with a trend towards convergence and integration. KETs

> can assist technology leaders in other fields to capitalise on their research effort (European Commission, 2018, p. 15-16)

The same source acknowledges that a much wider range of technology types may be relevant to consider strategic considerations, disruptive potential and/or relevance in relation to global grand challenges (European Commission, 2018, p. 20-22). We can differentiate this from our emerging concept of metatechnology in that it is construed in relation to strategic and geopolitical priorities rather than just in reference to the properties of a technology.

For the purposes of this paper, the main contentions are that some technologies may be qualitatively different from others based on the degree to which they enable the innovation of further technology; that this makes them particularly relevant from a responsible innovation perspective in terms of their ability to impact on society; and that this is particularly likely to be the case for digital technologies. To define metatechnology for our purposes, we can amalgamate the definitions of Braman and Glazer as follows: they are *core information technologies upon which others are based, and whose use vastly expands the degrees of freedom with which humans can act in the social and material worlds*.

THE CASE OF DIGITAL TECHNOLOGIES

Digital technologies (those using data in digital form) provide examples of emerging technologies which in many cases have metatechnological attributes relevant to considering alignment with societal needs. Brynolffson and McAfee (2014) provide a highly-cited case for the disruptive potential of digital technologies, with particular attention to their exponential and recombinant characteristics and zero marginal cost of reproduction.

The problematic aspects of digital technologies from a responsible innovation perspective are explored in detail by Jirotka *et al.* (2017), building on earlier work by Moor (1985, p. 269) and others and incorporating evidence from IT researchers and representative bodies. Their observations are summarised in Table 1.

Table 1. Problematic aspects of digital technologies for responsible innovation

Item	Description
Logical malleability and interpretive flexibility	Technology applications are often 'socially produced', and local innovations can result in unexpected uses
Prevalence and impact	Digital technologies increasingly shape labour markets and our daily lives
Pace	Compared to developments in the physical and life sciences, outputs may be developed, released and proliferate in a matter of hours
Difficulty predicting the uses of research outcomes	Researching objects in their contexts of use is often not possible and user adaptation can change the trajectory of digital technologies
Distributed development	Digital technology development is frequently split between different individuals, and often across many organisations[3].
Pacing problems	Impacts of technologies are increasingly only seen once they are in widespread use
Practical issues of embedding responsible innovation into professional responsibilities	It is difficult to define the relative roles of researchers and practitioners at the commercial interface and this requires collective action
Scope, complexity and convergence	The increasingly pervasive nature of technologies, often combined with rapid development, blurs boundaries between systems, features and functionality.

Source: summarised from Jirotka *et al.* (2017).

The problems of scope, pace and logical malleability are of hermeneutic interest – by the time the implications of a digital technology have been assessed, its use may have changed. This can be observed where companies provide APIs and SDKs (automatic programming interfaces and software development kits) to encourage integrations with their service, which can result in unexpected emergent uses of data as in the Facebook / Cambridge Analytica scandal (Berghel, 2018).

The increasing complexity of computational approaches brings new problems. While in some cases we can attribute these to the purposes and values of end users, there is evidence that algorithmic bias may be an intrinsic feature, rather than an avoidable design flaw of big data and machine learning-based approaches, or at the

3 This may involve international arbitrage, for example the coding of images using platforms such as Mechanical Turk by staff in low-income countries. The problem of responsibility attribution between developers and users in complex software development chains is discussed by Wolf *et al.* (2019).

very least is extremely difficult to 'design out' when bias is inherent in the social context of use (Beale *et al.,* 2020; Criado-Perez, 2019; Cheong *et al.*, 2021).

The issue is framed by de Reuver *et al.* (2020) who contrast the ontological uncertainty generated by digital technologies whose uses are determined by end users, with the more general epistemic uncertainty that exists at the design stage of other technologies. In this sense, digital technologies pose a qualitatively different problem, only partially soluble through steps such as broader and/or whole-lifecycle value-sensitive design approaches.

Digital technologies may also have upside implications for responsible innovation, in facilitating the exchange of ideas and open discussion, rendering database searches far more accessible and opening up new research methods (Bautista *et al.,* 2018). It is hard to see how, absent digital technology, academic efforts to research pandemic vaccines and the continuation of conferences and meetings could have taken place at the same rate.

While recent years have seen increased interest in ethical aspects of artificial intelligence and machine learning technologies from governments and organisations, methods of designing ethical concerns into systems are nascent, and regulation in this area chiefly consists of broad principles (Winfield *et al.,* 2019). Stahl *et al.* (2019, p. 376) similarly highlight "gaps in the fabric of responsibilities that govern ICTs".

In considering the metatechnological aspects of digital technologies we should also consider the extent to which they can originate from non-traditional modes of innovation, and may themselves dynamically transform networks of innovation (van de Poel, 2003). By implication, the effect of digital technologies in expanding the potential for different and potentially unexpected uses is potentially multiplicative and nonlinear. Some emerging digital technological trends have particular implications for the pace, complexity and scope of downstream development:

- The increasing tendency of software platforms to provide automatic programming interfaces and software development kits (APIs and SDKs) that allow for downstream development and integrations of services (Borgogno & Colangel, 2019)
- The open-source software movement, increasingly adopted by major software providers (Warren, 2020)
- The creation of low- and no-coding software development tools in general (Koksal 2019), and in particular low- and no-coding tools that allow non-experts to create and use machine learning models
- The increasing availability and scope of large datasets, in general and within organisations (George *et al.*, 2014)

- The exponential increase in internet-of-things connectivity (Nordrum, 2016)
- The development of new forms of digital manufacturing (e.g. Jensen-Haxel, 2011)
- Vertical integration of software platforms enabling the creation of more detailed datasets with the potential for more precise targeting, and limited state antitrust action (Kimmel & Kestenbaum, 2014)
- The growing tendency towards virtualisation and containerisation of software, enabling more rapid deployment and uptake (Silver, 2017).

Returning to the fundamental challenges for responsible innovation outlined by Blok and Lemmens (2015), it is apparent that the features of digital technologies in general, and of these emerging aspects in particular, pose specific problems associated with their 'metastatic' properties. Logical malleability and pace incur both epistemic insufficiency and ontological uncertainty. They are susceptible to differing interests among stakeholders leading to power imbalances, a particular issue in the case of increasingly prevalent machine learning approaches which are associated with algorithmic transparency (Hoadley *et al.*, 2010), and bias issues (e.g. Dastin, 2018), with approaches to transparency often constrained by commercial concerns, in Faustian business models whose nature is only belatedly beginning to be understood (Tibken, 2018).

One way to consider these issues is to suggest that digital technologies increase the 'RRI space' defined by Stahl (2013) based on their potential to significantly extend the range of actors, activities and societal norms that are potentially relevant to consider. In this sense, they will often constitute metatechnologies and as such are a relevant prism through which to assess the higher-order challenges metatechnologies may pose for responsible innovation discourse.

To consider the validity of foundational RRI accounts in relation to these challenges – or conversely, the extent to which they may have been developed with reference to issues associated with a limited range of technologies – we now assess the extent to which foundational RRI accounts have considered digital and metatechnologies and their associated problems.

METHODOLOGY

Critical hermeneutics enables investigation of the axiological and ontological assumptions in published accounts. The application of a hermeneutic perspective to Responsible (Research and) Innovation has been pioneered by Grunwald (2014, 2019, 2020), who draws attention to the importance of understanding the sometimes-contested meanings and technological futures attributed to new and emerging technologies. He identifies benefits of this perspective as avoiding epistemological over-caution, and as preferable to prognostic and scenario-based orientations in the case of 'overwhelming uncertainty'.

For our purposes, we can note that concern is particularly relevant in the case of technologies where there is limited evidence of impact, and high uncertainty over effects. The case of nanotechnology illustrates this – in the context of limited insights from early-stage research of a potentially metatechnological category of innovation, the meanings assigned to technologies came to dominate discussion (Simakova & Koenen, 2013; Fries, 2018).

The method used in this study adopts the approach of the hermeneutic study of RRI's foundational assumptions carried out by Timmermans and Blok (2018). In this case, rather than an inductive approach to discovering axiological assumptions of each account, a combined inductive and deductive approach, of analysing and categorising the technology examples referred to in each account will be used.

The rationale for a hermeneutic study is set out clearly by Timmermans and Blok (2018, p. 5). For the purposes of this study, key features are as follows:

- Critical hermeneutics is a tradition developed by Ricoeur (1981), Ricoeur and Thompson (1981), and Habermas (1978, 1988, 1990).
- It incorporates features of both the hermeneutic and critical theory traditions and aims to transcend taken-for-granted paradigms and critically examine their assumptions and practices.
- The position of the investigator relative to the phenomenon investigated should be considered.

The researcher perspective on this occasion is that of a small interdisciplinary team that includes academic interests in computing and social responsibility and a practitioner-researcher with experience introducing and overseeing the use of systems in organisations, including through contact with user groups and other organisations using third party software. This may be relevant in imbuing sensitivity

both to broader responsibility challenges of digital technologies and to the ways organisations and users can adapt and configure software.

Adapting the method of Timmermans and Blok (2018), the approach used here is as follows.

Source selection

The foundational accounts of RRI identified by Timmermans and Blok were adopted as the focus of enquiry. While other accounts of RRI exist, the validity of the selection of these accounts based on the criteria of a comprehensive, original, and influential framework or definition is reflected in the volume of citations the relative accounts have received in the period since publication (Loureiro & Conceição, 2019), and spans both the political, and academic perspectives we have noted can be identified within RRI (Owen, 2019).

For the purposes of this study, the text of accounts was defined as the following. With a view to validity of comparisons, the wordcount excluding references was assessed to contextualise any frequency-based observations:

- EC: European Commission (2012a, 2012b)
- VS: von Schomberg (2013)
- SOM: Stilgoe *et al.* (2013)
- VDH: van den Hoven (2013, 2017)

These sources reproduce those used by Timmermans and Blok (2018), with the exception that for the EC and VDH accounts an additional source is provided by the author which details technology examples considered in the main account.

Analysis of axiological and ontological assumptions per account

These sources were then subjected to critical hermeneutic analysis to identify implicit ontological assumptions. In this case, the assumptions of interest are the reference basis for each account in terms of the different examples of technological innovation they use, and the features of digital technology that are potentially problematic from a responsible innovation perspective they consider. In this sense, there is a focus on identifying and interpreting the examples in the text that illustrate the problems or issues that need to be addressed.

To relate assumptions of sources to the concept of metatechnology introduced above, a deductive coding approach was used. The documents were coded by two team members independently according to a pre-defined coding structure. The results of the coding were compared, any discrepancies discussed and clarified and a final decision made in order to ensure a common understanding was reached. The

coding structure was further refined during this process in light of emerging patterns. Table 2 summarises the protocol and definitions applied.

Table 2. Protocol and definitions

Step 1. Identify technology examples referenced by the study	Definition: 'An assemblage of practices and components'. In this sense the focus is on the particular innovation. Include any mention, count number of types not occurrences.
Step 2: Assess whether the technology example is an information technology (IT)	Definition: 'concerned with the dissemination, processing, and storage of information, esp. by means of computers' (Oxford English Dictionary). In particular, that the products are informational in nature or software.
Step 3: Assess whether the technology example is of a type identified in the academic literature as a metatechnology	Definition: the technology is one of the examples identified as a metatechnology in Braman (2004) or Jovanovich & Rousseau (2005). The list used is provided in the accompanying data table.
Step 4: Assess whether the technology shows characteristics of a metatechnology	Definition: technology matches all of: - 'is a core technology upon which others are or can be based' - 'vastly expands the degrees of freedom with which humans can act in the social and material worlds' - 'potential for high degree of reconstitution in use'
Step 5: Identify instances where the account discusses specific challenges of digital technologies	Definition: reference to any of the specific challenges itemised in Table 1
Step 6: Count instances where the account discuss challenges associated with onward innovation / reconstitution in use	Definition: discusses any features of technologies that increase the likelihood of it enabling further innovations

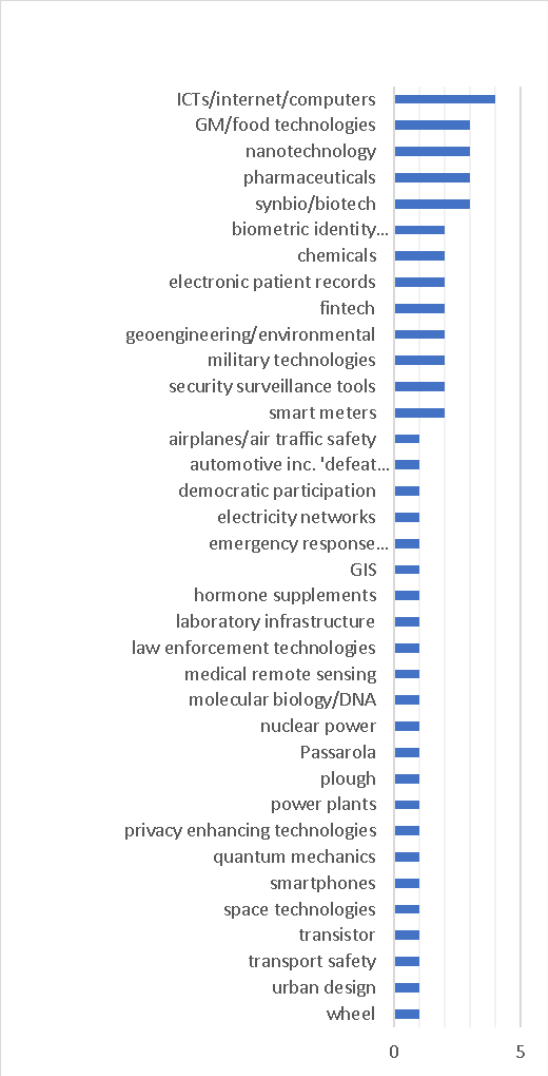
Source: Definitions gathered by the authors (Bryce *et al.*, 2022).

Coss-comparison of accounts

Based on the output of the previous stage, the results from each account were compared to enable a critical reflection on the scope of references.

Figures 1 and 2 show frequency of technology type by account. Table 3 summarises the examples returned by classification as technology, identified metatechnology type and metatechnology characteristic matches, and number of references to digital technology and reconstitution in use issues.

Figure 1. Technologies cited across all RRI accounts by frequency (n=55)



Source: Data processed by the authors (Bryce *et al*. 2022).

Figure 2. Technology categories referenced by account (n=55)

	EC	VS	SOM	VDH
airplanes/air traffic safety				■
automotive inc. 'defeat devices'				■
biometric identity management	■	■		
chemicals		■	■	
democratic participation		■		
electricity networks				■
electronic patient records		■		■
emergency response technologie	■			
fintech			■	■
geoengineering/environmental			■	■
GIS				■
GM/food technologies	■	■	■	
hormone supplements		■		
ICTs/internet/computers	■	■	■	■
laboratory infrastructure	■			
law enforcement technologies	■			
medical remote sensing		■		
military technologies		■		■
molecular biology/DNA			■	
nanotechnology	■	■	■	
nuclear power				■
Passarola		■		
pharmaceuticals	■	■	■	
plough				■
power plants				■
privacy enhancing technologies	■			
Quantum mechanics			■	
security surveillance tools	■	■		
smart meters		■		■
smartphones				■
space technologies		■		
synbio/biotech	■	■	■	
transistor				■
transport safety				■
urban design				■
wheel				■

Source: Data processed by the authors (Bryce *et al*., 2022).

Table 3. Technology examples cited in RRI accounts by type

Account	n[4]	Wordcount	IT examples	Metatechnology type examples[5]	Metatechnology characteristic examples[6]	Digital problem occurrences	Reconstitution in use problem occurrences
EC	11	22737	6	7	4	10	0
VS	16	10100	6	7	4	4	1
SOM	10	10145	2	4	6	0	0
VDH	18	13034	5	6	3	6	0

Source: Data gathered by the authors (Bryce *et al.*, 2022).

Table 4. Digital technology challenges cited in RRI accounts

	EC	VS	SOM	VDH
Logical malleability and interpretive flexibility	■	■		■
Prevalence and impact	■			
Pace	■			
Difficulty predicting the uses of research outcomes				
Distributed development	■	■		
Pacing problems	■			
Practical issues of embedding responsible innovation into professional responsibilities	■			■
Scope, complexity and convergence	■			■
Reconstitution in use / Onward innovation		■		

Source: Data processed by the authors (Bryce *et al.*, 2022).

4 Number of technology types cited in the account.

5 Proportion of examples that are an identified metatechnology type – see Table 2.

6 Proportion of examples assessed as having metatechnology characteristics – see Table 2.

ANALYSIS

European Commission

Overview

The European Commission account of RRI was set out in a policy statement in 2012 (European Commission, 2012a), and revisited in later publications and declarations before being integrated into funding calls. While Timmermans and Blok (2018) note that RRI literature recognise these accounts as an authentic source of the EC account on RRI, the 2012 statement differs from other accounts in that it asserts a policy agenda – it is a policy document, rather than an academic paper. To enable a comparison on the basis of underpinning technology examples used in constructing the account, the 'Science with and for Society' (SwafS) report on 'Ethical and regulatory challenges to Science and Research Policy at the Global Level' (European Commission, 2012b) presented by the SwafS expert group with EU and US membership is assessed as a source document informing the development of the EC account.

Characterisation of technology examples

A relatively high proportion of examples in the EC source material were classified as having attributes of metatechnologies, mainly because information technologies featured prominently as examples in the source material analysed. The account touches on nearly all the challenges of digital technology identified by Jirotka *et al.* (2017) while adding another (inherent transparency of digital data), although challenges of onward innovation are not discussed.

The prevalence of digital examples may reflect salient political issues for the EC during the account's development in 2010-2012, in particular the development of the the General Data Protection Regulation through the EC data protection reform package combined with the introduction of the EuroSur border surveillance programme and accompanying. While not discussed within the reference documents, the emergence of privacy and data protection concerns associated with EuroSur might otherwise have been developed as an example of the need to anticipate impacts associated with the ability of digital technologies to be adapted for alternative uses (in this case, surveillance beyond that necessary for border security – Marin, 2011).

Von Schomberg

Overview

The von Schomberg account, presented over a series of studies is widely cited in Responsible Innovation literature and is referenced in the Stilgoe, Owen and Macnaghten account. Written during the author's tenure as an official of the European Union, the account has similarities to the EC account but is explicitly a personal rather than institutional vision. The account's emphasis on redefining the 'right impacts of innovation' in broader societal rather than macro-economic terms is substantiated through several examples. In line with Timmermans and Blok (2018), von Schomberg (2013) is recognised as the authoritative account.

Characterisation of technology examples

The VS account provides a broad vista of examples, including digital technologies, offering the largest number in a single study and exceeded only by the VDH account when the latter is considered to include an additional reference paper. Together with the EC account it has the highest number of examples identified as metatechnologies, including discussions on nanotechnology and synthetic biology as well as information technologies, and uniquely among the accounts were the potential of digital technologies to impact democratic participation is considered (p. 7-8), along with two mentions of digital technology challenges the issue of onward innovation: via a discussion of adapted uses of Microsoft Kinect systems, the importance from a responsible innovation perspective of considering the potential for recombinant uses (p. 7), and a discussion on the need for responsible innovation to consider distributed development issues (p. 13) rather than a sole focus on individual responsibility. This latter issue has recently been expanded on in the context of the discourse in responsible innovation in digital technologies, with reference to ecosystems approaches (e.g. Stahl, 2021).

Stilgoe, Owen and Macnaghten

Overview

The account of Stilgoe, Owen and Macnaghten is frequently cited in RRI literature and while drawing on and elaborated in a wider series of papers is broadly seen (e.g. Timmermans & Blok, 2018) as embodied in Stilgoe *et al.* (2013). This paper provides a broad overview of issues and emergent methods in relation to responsible innovation, categorising these in a way which has been widely adopted, particularly in the United Kingdom through the Engineering and Physical Sciences Research Council. There is arguably a difference in emphasis within the account compared to the others on a review of methods, and the use of an in-depth case study, rather than an attempt to

illustrate through breadth of examples. The SPICE – Stratospheric Particle Injection for Climate Engineering geoengineering project is used to illustrate the potential for responsible innovation methods to alter technological trajectories, but for the purpose of our analysis is not categorise as a digital or metatechnological example.

Characterisation of technology examples

The SOM account is superficially similar to the EC and VS accounts in the scope of technology examples referred to, but other than brief discussions the focus is on geoengineering as the primary case and no references to the specific challenges of digital technology or onward innovation issues were identified.

Van den Hoven

Overview

The account of Jeroen van den Hoven synthesises a body of work on value-sensitive design (VSD). It is relevant to this study that the approach, which aims to expose and integrate values into design processes originates from studies of information technologies. The approach was set out in van den Hoven (2013), and further studies have strengthened the relevance of VSD to responsible innovation debates (for example de Reuver *et al.*, 2020).

With a view to equivalence of wordcount, the van den Hoven corpus studies for the purposes of analysis was extended to include van den Hoven (2017). This text, while published later, expands on the technological reference points for the theory of VSD-based responsible innovation and is cited as a work in progress by the main account (van de Hoven, 2013), so as with the EC account is assessed as being part of the reference text.

Characterisation of technology examples

The *point de depart* for the VDH account is digital technologies, and the text begins by focussing on the Netherlands electronic patient record and smart meter programme sagas. A wide range of technological examples are cited, with a tendency towards physical engineering disciplines but several discussions of software and human-computer interface aspects are also included. No reference to onwards innovation challenges was identified. While reference to issues such as the contested introduction of smart metering into the Netherlands usefully expose complexity challenges associated with digital technologies, the central thesis – that design teams should actively consider the values they are applying to their development – is liable to challenge in the case of technologies which, once introduced, may be relatively freely appropriated by different actors and which may not realistically be constrained to uses associated with a developer's explicitly intended values.

DISCUSSION

The intent of this paper is to critically examine the underpinnings of paradigmatic accounts of R(R)I, in particular with reference to digital technologies and the concept of metatechnologies, and to consider the broader implications for responsible innovation if these accounts have been constructed using examples that may not consider certain types of technology.

There are two clear considerations in our findings; the first, that, insofar as there is any suggestion that RRI accounts may not have included digital and metatechnologies – and that the discourse may be in need of revision based on this – the findings do not on first inspection support this conclusion. Each foundational account includes digital technologies within its technological references, and also includes references to technologies that could be considered metatechnologies. This suggests that the discourse on RRI, while to an extent grounded within the physical sciences through the prominence of the SOM account – is not inherently limited in its consideration of innovation from a digital technological perspective. This offers scope for the discourse on RRI to continue to influence responsible innovation as practice, as many large scale research projects focus on digital innovation (European Commission, 2021). However, simply identifying digital technologies within the accounts offers only a limited perspective on the realities of the representation of digital technologies, and metatechnologies in particular, within the RRI discourse.

Therefore, secondly, it is worth considering the extent to which the challenges of digital technologies for RRI are considered within the accounts. In many cases references to digital technologies are brief or superficial; whilst a variety of digital technologies are identified, the discussion around the challenges of these technologies is predominantly limited, and occasionally absent entirely. Here, then, the roots of responsible innovation discourse in physical science disciplines show more clearly, as even the accounts that do offer some consideration of challenges in relation to digital technologies, do not generally consider all responsible innovation challenges of digital technology. In particular, the accounts do not (other than the account of von Schomberg) assess issues of distributed development, or onward innovation issues associated with technologies that can be reconstituted in use. The SOM account – potentially the most influential, per Loureiro and Conceição (2019) – is developed through a physical sciences case and does not discuss challenges associated with digital technology or others identified as metatechnologies. This suggests that some problematic features of digital and/or metatechnologies have not been fully considered in foundational RRI accounts.

The overlap in technology examples considered by the accounts (Figure 2 above) highlights the collaborative spirit and contemporaneous timeframe within which the core works of RRI were produced. This may be a source of reassurance in that accounts largely agree in the technology scope they feel is appropriate to illustrate the concept of RRI – but it may also be an indication that just as the discipline of Science and Technology Studies was developed largely in response to specific nanotechnology concerns, Responsible Research and Innovation may at least to an extent be founded on twentieth and early twenty-first century technological problems, with the implication that the methodologies it prescribes may become less relevant for emerging digital technologies.

As theories develop in a particular historic context, it is reasonable to suggest that a restatement of RRI accounts in the 2020s might feature challenges of digital technology assessment and governance more prominently. Equally, this is a challenge – to maintain the constantly self-critical aspect called for by foundational RRI authors, there is a need to revisit the problems that RRI should address, and the techniques necessary to address them, vis-à-vis digital technology. The particular difficulties of applying RRI principles to the development of digital technologies is clearly understood, and some recommendations are beginning to emerge in an attempt to develop RRI into a framework capable of addressing these difficulties alongside promoting socially desirable innovation (Jirotka *et al.*, 2017). However, given the situated nature of the development of RRI within a particular historical-technological context, questions should be asked as to whether a fundamental shift towards ever more prevalent digital metatechnologies might require a reconsideration of RRI itself; can a responsible innovation approach developed to address social and ethical issues in the physical sciences adequately translate to the social and ethical issues of emerging, disruptive digital and metatechnologies? Whilst the answer to this question is still contested, it is notable that doteveryone, a Non-Governmental Organisation, proposed a shift towards a 'responsible by design' (Miller & Ohrvik-Stott, 2018) approach on the basis of specific social and ethical issues in relation to digital technologies, incorporating, for example, fundamental human rights that may be covertly elided by digital (meta)technologies that are cross-cutting in nature.

This study therefore provides general support to the validity of seminal RRI accounts, while highlighting the need for further analysis of new and emerging digital technologies and in particular their capacity for enabling onward innovation. In this sense, Braman's (2007) sense of the degree to which a technology can shape or create new information flows may be most relevant. The concept of 'metatechnology' is in this sense a question of degree, rather than of type – the concepts of 'barriers to entry' and 'marginal cost of reproduction' may provide useful measures of technologies' potential to proliferate and 'mutate', for technology assessment

purposes. This may surface the implications of inventions that democratise innovation or enable users to recombine disparate datasets and services such as APIs and low-coding tools, the network-level potential of connected devices at scale, and assumptions inherent in, for example, the movement towards open-source software development.

At the same time, it invites us to consider, in the face of increasingly ubiquitous digital technology, the extent to which the entire apparatus of RRI may still need to be reconsidered. If the notion that responsible innovation implications can be anticipated at an early stage is a central tenet of RRI, but digital technologies increasingly enable an exponentially wider range of applications – and are developed in a distributed fashion such that teams working on components may not be able be aware of their broader implications – we may increasingly need to redefine the locus of responsible innovation further downstream, in the organisations and individuals who configure the use of logically malleable digital technologies. This may be particularly true if the diffusion of digital metatechnologies accelerates the rate at which technological innovation takes place without the input of the scientific community. (Godin, 2016)

This may in turn have implications for the regulation of potential metatechnologies (for example 3D printing), and supports arguments that where uses and societal impacts cannot reliably be anticipated, innovation policy and responsible innovation assessments should increasingly consider prevention, or exnovation of technologies as a valid and potentially desirable outcome. In this sense, our findings support the argument by Owen and Pansera (2019) that for responsible innovation activities to apply meaningfully to the broader innovation ecosystem, given the environmentally, politically and ethically entangled of disruptive technovisionary innovations, it will increasingly be necessary to consider political as well as technical dimensions of governance.

It is also worth reflecting that RRI, in particular, is specifically situated with the European political apparatus (Owen, 2019), and as such may presume aspects of technology use, innovation mechanisms, or responsible innovation practices that are region-specific (Wakunuma *et al.*, 2021). Technologies subject to strict regulation in the West such as facial recognition may not be similarly constrained in other cultures with the result that they proliferate, mature, and develop new applications with global implications. Similarly, low barriers to entry in one region may be insurmountable in others, for example due to issues around access to broadband and computing power, and this may constrain onward innovation. As such, the increasing prominence of digital (meta)technologies also implies the need for an increased focus on non-Western innovation systems. This prospect alone raises questions about the feasibility

of the translation of even the fundamental concepts of responsible innovation with respect to technologies with deeply diverse impacts across and between global regions.

CONCLUSION

While foundational accounts of RRI do include references to digital and metatechnologies, the brevity of these references and the possibility that they do not consider all specific challenges associated with these technologies mean that further exploration and theorisation of responsible innovation in relation to digital technologies is required to maintain the relevance of responsible innovation disciplines in the face of emerging technologies and practices. These may need, in particular, to consider the way in which they enable onwards innovation in different cultural and organisational contexts, and to continuously seek clarification of the futures specific technological developments may enable, either themselves or in combination with other emerging technologies.

As a final reflection, Stilgoe *et al.* (2013, p. 32) invoke a viral analogy in their suggestion that emerging technologies pose additional challenges to governments (and by extension, organisations) in the sense that they encounter organisms to which they have not yet developed a regulatory 'immune response'. This may point to new research directions drawing on evolutionary economics methods to develop a 'genetics of technology', but more significantly, it suggests that responsible innovation's potential to inoculate society from technology harms and connect innovations to the 'right ends' is dependent on our awareness of technology's state of the art, and the innovatory vectors through which it develops.

Thus, while this study has found that paradigmatic narratives of responsible research and innovation do not neglect digital technologies and those we can identify as metatechnologies, the continuing value of RRI as a discourse to our society will depend on researchers' and practitioners' detailed awareness of the potential of these technologies for cascading, downstream innovation.

REFERENCES

Abernathy, W. J., & Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction. *Research Policy*, 14(1), 3-22. https://doi.org/10.1016/0048-7333(85)90021-6

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves* (1. Free Press pbk. ed). Free Press.

Bautista, S., Mazaj, J., & Cárdenas, M. (2018). Developing RRI Practices: The Role of the ICT in Stakeholders' Knowledge Exchange and Co-creation Processes. In F. Ferri, N. Dwyer, S. Raicevich, P. Grifoni, H. Altiok, H. T. Andersen, Y. Laouris, & C. Silvestri, *Responsible Research and Innovation Actions in Science Education, Gender and Ethics* (pp. 21-26). Springer International Publishing. https://doi.org/10.1007/978-3-319-73207-7_4

Beale, N., Battey, H., Davison, A. C., & MacKay, R. S. (2020). An unethical optimization principle. *Royal Society Open Science*, 7, 200462. https://doi.org/10.1098/rsos.200462

Berghel, H. (2018). Malice Domestic: The Cambridge Analytica Dystopia. *Computer*, 51(5), 84-89. https://doi.org/10.1109/MC.2018.2381135

Bessant, J. (2013). Innovation in the Twenty-First Century. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 1-25). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch1

Blok, V. (2020). What is Innovation? Laying the Ground for a Philosophy of Innovation. *Techné: Research in Philosophy and Technology*, 25(1), 72-96. https://doi.org/10.5840/techne2020109129

Blok, V., & Lemmens, P. (2015). The Emerging Concept of Responsible Innovation. Three Reasons Why It Is Questionable and Calls for a Radical Transformation of the Concept of Innovation. 19-35. https://doi.org/10.1007/978-3-319-17308-5_2

Borgogno, O., & Colangelo, G. (2019). Data sharing and interoperability: Fostering innovation and competition through APIs. *Computer Law & Security Review*, 35(5), 105314. https://doi.org/10.1016/j.clsr.2019.03.008

Braman, S. (Ed.). (2004). The Meta-Technologies of Information. In *Biotechnology and communication: The meta-technologies of information* (pp. 3-38). Lawrence Erlbaum Associates.

Bross, I. D. J. (1981). Metatechnology: A technology for the safe, effective, and economical use of technology. *Metamedicine*, 2(2), 145-153. https://doi.org/10.1007/BF00884419

Cressman, D. (2019). Disruptive Innovation and the Idea of Technology. *Novation-Critical Studies of Innovation*, 1(June), 18-40. http://www.novation.inrs.ca/index.php/novation/article/view/7

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies (First Edition). W. W. Norton & Company.

Cheong, M., Lederman, R., McLoughney, A., Njoto, S., Ruppanner, L., & Wirth, A. (2021). *Ethical Implications of AI Bias as a Result of Workforce Gender Imbalance*. University of Melbourne. https://about.unimelb.edu.au/__data/assets/pdf_file/0024/186252/NEW-RESEARCH-REPORT-Ethical-Implications-of-AI-Bias-as-a-Result-of-Workforce-Gender-Imbalance-UniMelb,-UniBank.pdf

Christensen, C. M. (1997). The innovator's dilemma: When new technologies cause great firms to fail. Harvard Business School Press.

Christensen, C., Raynor, M., & McDonald, R. (2015). What Is Disruptive Innovation? *Harvard Business Review*, December, 44-53.

Criado-Perez, C. (2020). Invisible women: Data bias in a world designed for men. Abrams.

Culot, G., Nassimbeni, G., Orzes, G., & Sartor, M. (2020). Behind the definition of Industry 4.0: Analysis and open questions. *International Journal of Production Economics*, 226, 107617. https://doi.org/10.1016/j.ijpe.2020.107617

Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters, October 10. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

David, P. A., & Wright, G. (2006). The Economic Future in Historical Perspective. British Academy.

de Reuver, M., van Wynsberghe, A., Janssen, M., & van de Poel, I. (2020). Digital platforms and responsible innovation: Expanding value sensitive design to overcome ontological uncertainty. *Ethics and Information Technology*, 22(3), 257-267. https://doi.org/10.1007/s10676-020-09537-z

European Commission. (2012a). *Responsible Research and Innovation – Europe's ability to respond to societal challenges (KI-31-12-921-EN-C)*. https://op.europa.eu/en/publication-detail/-/publication/2be36f74-b490-409e-bb60-12fd438100fe

European Commission. (2012b). *Ethical and Regulatory Challenges to Science and Research Policy at the Global Level (KI-NA-25-101-EN-C)*. https://op.europa.eu/en/publication-detail/-/publication/84fc3de4-6641-4d9e-be58-9ca7da3d397b/language-en

European Commission. (2018). *Re-Finding Industry*. https://op.europa.eu/en/publication-detail/-/publication/28e1c485-476a-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-116043022

European Commission. (2021). *Horizon Europe Strategic Plan 2021-2024*. https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/3c6ffd74-8ac3-11eb-b85c-01aa75ed71a1

Feenberg, A. (2017). *Technosystem: The social life of reason*. Harvard University Press.

Fries, M.-H. (2018). Nanotechnology and the Gray Goo Scenario: Narratives of Doom? *ILCEA*, 31. https://doi.org/10.4000/ilcea.4687

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326. https://doi.org/10.5465/amj.2014.4002

Gillin, P. (2020). Low-code and no-code tools may finally usher in the era of 'citizen developers'. *Silicon Angle*, October 6. https://siliconangle.com/2020/10/06/low-code-no-code-tools-may-finally-usher-era-citizen-developers/

Glazer, R. (2007). Meta-Technologies and Innovation Leadership: Why There May Be Nothing New under the Sun. *California Management Review*, 50(1), 120-143. https://doi.org/10.2307/41166419

Godin, B. (2016). Technological Innovation: On the Origins and Development of an Inclusive Concept. *Technology and Culture*, 57(3), 527-556. https://doi.org/10.1353/tech.2016.0070

Grunwald, A. (2014). The hermeneutic side of responsible research and innovation. *Journal of Responsible Innovation*, 1(3), 274-291. https://doi.org/10.1080/23299460.2014.968437

Grunwald, A. (2019). Responsible innovation in emerging technological practices. In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation. A Global Resource* (pp. 326-338). Edward Elgar Publishing. https://doi.org/10.4337/9781784718862.00031

Grunwald, A. (2020). The objects of technology assessment. Hermeneutic extension of consequentialist reasoning. *Journal of Responsible Innovation*, 7(1), 96-112. https://doi.org/10.1080/23299460.2019.1647086

Hoadley, C. M., Xu, H., Lee, J. J., & Rosson, M. B. (2010). Privacy as information access and illusory control: The case of the Facebook News Feed privacy outcry.

Electronic Commerce Research and Applications, 9(1), 50-60. https://doi.org/10.1016/j.elerap.2009.05.001

Hogendorn, C., & Frischmann, B. (2020). Infrastructure and general purpose technologies: A technology flow framework. *European Journal of Law and Economics*, 50(3), 469-488. https://doi.org/10.1007/s10657-020-09642-w

Jirotka, M., Grimpe, B., Stahl, B., Eden, G., & Hartswood, M. (2017). Responsible research and innovation in the digital age. *Communications of the ACM*, 60(5), 62-68. https://doi.org/10.1145/3064940

Jonas, Hans (1979). Toward a Philosophy of Technology. *Hastings Center Report*, 9(1), 34-43.

Jovanovic, B., & Rousseau, P. L. (2005). General Purpose Technologies. In *Handbook of Economic Growth* (Vol. 1, pp. 1181-1224). Elsevier. https://doi.org/10.1016/S1574-0684(05)01018-X

Kagermann, H., Wolf-Dieter, L., & Walster, W. (2011). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. *VDI Nachtrichten*, 13(2). https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/News_Media/Presse/Presse-Highlights/vdinach2011a13-ind4.0-Internet-Dinge.pdf

Kendall, K. E. (1997). The Significance of Information Systems Research on Emerging Technologies: Seven Information Technologies that Promise to Improve Managerial Effectiveness. *Decision Sciences*, 28(4), 775-792. https://doi.org/10.1111/j.1540-5915.1997.tb01331.x

Kimmel, L., & Kestenbaum, J. (2014). What's Up with WhatsApp? A Transatlantic View on Privacy and Merger Enforcement in Digital Markets. *Antitrust*, 29, 48-54.

Koksal, I. (2019). The Rise Of Low-Code App Development. *Forbes*. https://www.forbes.com/sites/ilkerkoksal/2020/04/29/the-rise-of-low-code-app-development/?sh=4fc5d8186807

Loureiro, P. M., & Conceição, C. P. (2019). Emerging patterns in the academic literature on responsible research and innovation. *Technology in Society*, 58. https://doi.org/10.1016/j.techsoc.2019.101148

Marin, L. (2011). Is Europe Turning into a 'Technological Fortress'? Innovation and Technology for the Management of EU's External Borders: Reflections on FRONTEX and EUROSUR. In M. A. Heldeweg & E. Kica (Eds.), *Regulating Technological Innovation* (pp. 131-151). Palgrave Macmillan UK. https://doi.org/10.1057/9780230367456_8

Meyrowitz, J. (1995). Mediating Communication: What Happens? In J. Downing, A. Mohammadi, A. Sreberney-Mohammadi (Eds.), *Questioning the Media: A Critical Introduction*. Sage Publications

Mitcham, C. (1995). Notes toward a Philosophy of Meta-Technology. *Society for Philosophy and Technology Quarterly Electronic Journal*, 1(1), 13-17. https://doi.org/10.5840/techne199511/25

Miller, C., & Ohrvik-Stott, J. (2018). *Regulating for Responsible Technology—Capacity, Evidence and Redress: A New System for a Fairer Future*. Doteveryone. https://doteveryone.org.uk/wp-content/uploads/2018/10/Doteveryone-Regulating-for-Responsible-Tech-Report.pdf

Moor, J. H. (1985). What is Computer Ethics? *Metaphilosophy*, 16(4), 266-275. https://doi.org/10.1111/j.1467-9973.1985.tb00173.x

Nordrum, A. (2016). *Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated*. IEEE Spectrum. https://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated

Owen, R., & Pansera, M. (2019). Responsible Innovation: Process and Politics. In R. von Schomberg, & J. Hankins (Eds.), *International Handbook on Responsible Innovation: A Global Resource* (pp. 35-48). Elgar Publishing.

Owen, Richard. (2019). Responsible Innovation and Responsible Research and Innovation. In D. Simon, S. Kuhlmann, J. Stamm, & W. Canzler (Eds.), *Handbook on Science and Public Policy* (pp. 26-48). Edward Elgar Publishing. https://doi.org/10.4337/9781784715946.00010

Ribeiro, B. E., Smith, R. D. J., & Millar, K. (2017). A Mobilising Concept? Unpacking Academic Representations of Responsible Research and Innovation. *Science and Engineering Ethics*, 23, 81–103. https://doi.org/10.1007/s11948-016-9761-6

Rip, A., & Voß, J.-P. (2013). Umbrella Terms as a Conduit in The Governance of Emerging Science and Technology. *Science, Technology & Innovation Studies*, 9(2), 39-60.

Silver, A. (2017). Software simplified. *Nature*, 546(7656). https://www.nature.com/news/software-simplified-1.22059

Simakova, E., & Coenen, C. (2013). Visions, Hype, and Expectations: A Place for Responsibility. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 241-267). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch13

Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, 40(6), 708-716. https://doi.org/10.1093/scipol/sct067

Stahl, B. C. (2021). *Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies*. Springer. https://doi.org/10.1007/978-3-030-69978-9

Stahl, B. C., Borsella, E., Porcari, A., & Mantovani, E. (2019). Responsible innovation in ICT: Challenges for industry. In R. von Schomberg, & J. Hankins (Eds.), *International Handbook on Responsible Innovation. A Global Resource* (pp. 367-378). Edward Elgar Publishing. https://doi.org/10.4337/9781784718862.00034

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568-1580. https://doi.org/10.1016/j.respol.2013.05.008

Timmermans, J., & Blok, V. (2018). A critical hermeneutic reflection on the paradigm-level assumptions underlying responsible innovation. *Synthese*, 198, 4635-4666. https://doi.org/10.1007/s11229-018-1839-z

Vallenilla, E. M. (1999). From Meta-Technology to Ecology. *Bulletin of Science, Technology & Society*, 19(5), 411-415. https://doi.org/10.1177/027046769901900509

Van De Poel, I. (2003). The transformation of technological regimes. *Research Policy*, 32(1), 49-68. https://doi.org/10.1016/S0048-7333(01)00195-0

van den Hoven, J. (2013). Value Sensitive Design and Responsible Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 75-83). John Wiley & Sons Inc.

van den Hoven, J. (2017). The Design Turn in Applied Ethics. In J. van den Hoven, S. Miller, & T. Pogge (Eds.), *Designing in Ethics* (1st ed., pp. 11-31). Cambridge University Press. https://doi.org/10.1017/9780511844317.002

von Schomberg, L., & Blok, V. (2019). Technology in the Age of Innovation: Responsible Innovation as a New Subdomain Within the Philosophy of Technology. *Philosophy & Technology*, 34, 309-323. https://doi.org/10.1007/s13347-019-00386-3

von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation: Managing the*

Responsible Emergence of Science and Innovation in Society (pp. 51-74). John Wiley & Sons Inc. https://doi.org/10.1002/9781118551424.ch3

Wakunuma, K., de Castro, F., Jiya, T., Inigo, E. A., Blok, V. & Bryce, V. (2021). Reconceptualising responsible research and innovation from a Global South perspective. *Journal of Responsible Innovation*, 8(2), 267-291. https://doi.org/10.1080/23299460.2021.1944736.

Winfield, A. F., Michael, K., Pitt, J., & Evers, V. (2019). Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems [Scanning the Issue]. *Proceedings of the IEEE*, 107(3), 509-517. https://doi.org/10.1109/JPROC.2019.2900622

Wolf, M. J., Miller, K. W., & Grodzinsky, F. S. (2019). On the responsibility for uses of downstream software. In D. E. Wittkower (Ed.), *Computer Ethics – Philosophical Enquiry (CEPE) Proceedings* (14 pp.). https://doi.org/10.25884/7576-WD27

Wright, G. (2000). Book Reviews – General Purpose Technologies and Economic Growth. *Journal of Economic Literature*, 38(1), 161-162. https://doi.org/10.1257/jel.38.1.115

Between sustainability commitments and anticipated market requirements. Exploring the resilience of the techno-economic innovation paradigm in the midstream of construction research[1]

Deniz Frost*, Kathrin Braun and Cordula Kropp*****

* *University of Stuttgart*

** *University of Stuttgart*

*** *University of Stuttgart*

ABSTRACT

This article studies ways of dealing with the tension between a commitment to sustainable and responsible research and anticipated market requirements in the midstream of a research process in architecture and construction. Using a slightly modified version of Socio-Technical Integration Research (STIR), we explored the chances of questioning the primacy of the techno-economic innovation paradigm by deliberately provoking reflections through STIR interactions. Our research underlines the difficulties and limitations of challenging an orientation towards values of efficiency and productivity in favour of social and environmental values in the midstream of the research process and examines *how* the techno-economic innovation paradigm is able to insulate itself against critical questioning. It sheds light at the critical role of the underlying assumption that marketability of prospective outcomes is not one objective amongst others but the precondition for all others and at two argumentative patterns we termed the "lack-of-agency" and the "reconciliation-after-all" pattern.

Keywords: Socio-Technical Integration Research (STIR); Anticipated Market Requirements; Construction Industry; Responsible Research and Innovation (RRI); Constructive Technology Assessment.

Proposal Submitted 28 August 2020, Article Received 9 April 2021, Reviews Delivered 2 July and 7 October 2021, Revised 28 October 2021, Accepted 1 December 2021, Available online 28 February 2022.

[1] Acknowledgement: The research published in this article is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC 2120/1–390831618. The authors extend their gratitude to the DFG. Furthermore, the authors wish to thank Erik Fisher for immensely helpful comments and advice and not least the two researchers who so generously participated in our STIR exercises and shared their thoughts, experiences and questions with us.

INTRODUCTION: The Need for Responsible Innovation in Architecture and Construction

The construction industry is in need of social, cultural and technological innovation. At present, the industry is a major contributor of greenhouse gas emissions; in 2019, energy-related CO_2 emissions from building operations and construction reached their highest recorded level and accounted for 38 percent of total global energy-related CO_2 emissions (UNEP, 2020). In addition, the industry is responsible for excessive use of raw materials and is a major producer of waste. According to OECD estimates, the global construction sector will more than double from 2017 to 2060 and its use of materials will increase to almost 84 gigatons of construction materials in 2060 (OECD, 2019). Currently, 40 to 50 percent of the non-energy resources extracted for global materials are used for housing, construction and infrastructure (UNEP, 2020); in the European Union, construction and demolition waste accounts for approximately 25-30 percent of the total waste generated. Most of this, including concrete, gypsum, ceramics, metals, plastic, solvents, asbestos and excavated soil, is currently downcycled rather than recycled (UNEP, 2020). Furthermore, even additional construction of more sustainable houses would add to the problems of land use and land sealing.

At the same time, the demand for adequate housing continues to increase worldwide. The UN estimate that the world's human population will grow from 7.7 billion in 2019 to 9.7 billion in 2050 (UN, 2019), with more than two-thirds living in urban areas. Affordable and adequate housing is already seriously lacking. According to UN estimates, 1.8 billion people live in inadequate housing in slums or overcrowded settlements or in a state of homelessness (UN, 2020), exposed to global health crises such as COVID-19 and climate change-induced emergencies. In short, construction faces the double challenge of performing the transition towards a sustainable, net-zero emissions and zero-waste building culture and simultaneously delivering adequate, healthy and equitable housing for a growing world population.

From an economic perspective, construction is suffering mainly from an innovation, profitability and productivity crisis, allegedly due to its notorious aversion to innovation (Roland Berger, 2016; Ribeirinho *et al.*, 2020).

Governments invest their hopes in digital technologies to solve construction's multiple crises and are advocating for building information modelling (BIM) to become standard for public construction projects (Lee & Borrmann, 2020). While some actors voice concerns about job losses, de-skilling, a decline in architectural quality, and economic concentration processes, the dominant view is that digitalisation offers solutions to the housing crisis, the environmental crisis, the economic crises, the productivity crisis, and more recently also the COVID-19 crisis (Braun & Kropp, 2021).

An examination of the socio-technical visions and imaginaries underlying current debates on the digital transformation of architecture and construction (Braun & Kropp, 2021) shows a widespread agreement among industry, policy and civil society actors about the challenges and problems described above as well as nearly universal expectations that digital technologies will offer solutions to all of them. Possible conflicts between technological innovation and other objectives, such as a zero-emissions and zero-waste building culture or a liveable, equitable built environment, are rarely addressed in these discussions; the prevalent assumption is that digital transformation will automatically generate more sustainable, high-quality, socially adequate and acceptable buildings and construction processes. The latter are largely considered by-products of technologically conceived innovation. Borrowing Morozov's (2013) term, solutionism, we can see here a macro-level type of techno-solutionism, a belief that complex social problems can be ascribed to a lack of technological efficiency and process optimisation. Joly and Rip (2012) posit the "cornucopia" conception of technoscience, according to which technoscientific innovations would solve many of humanity's major problems if only sufficient resources could be mobilised to push them forward. Critical research into the interrelations between digitalisation and sustainability, however, has shown that solutionist and cornucopia conceptions do not hold; rebound effects and increased use of energy for servers and ICT operations, among other things, damage the environment to an extent that may outweigh the environmental benefits of digitalisation (Coroamă & Mattern, 2019; Lange *et al.*, 2020). To date, detailed research into the relationships between sustainability and digitalisation in the field of architecture and construction remains scarce (Zhang *et al.*, 2020).

TECHNO-SOLUTIONISM *VERSUS* RRI

There is no problem with technological solutions or technological efficiency per se; it is certainly in the common interest to develop efficient solutions for a sustainable, liveable and equitable built environment. Yet, from a social and environmental justice point of view, problems arise when technological efficiency is defined and measured first and foremost in terms of cost savings and profitability and the values of sustainability, fairness and justice are considered innovative only when they contribute to the former. Under the conditions of global competition, the authority to decide what qualifies as an innovative solution ultimately rests with the market. There is good reason, however, to agree with von Schomberg and Hankins (2019, p. 1) that "market innovations do not automatically deliver on societally desirable objectives". This concern gave rise to the paradigm of "responsible research and innovation" (RRI)

(von Schomberg, 2008; Owen *et al.*, 2012; Owen *et al.*, 2013; Burget *et al.*, 2017), defined by von Schomberg as...

> […] a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society). (von Schomberg, 2008, p. 50)

RRI is intended as a strategy to challenge the dominant techno-economic paradigm in research and innovation and counterbalance it with an orientation towards socially desirable objectives and solutions for the grand challenges laid down inter alia in the UN sustainable development goals (SDGs) (von Schomberg & Hankins, 2019). RRI, however, has also been criticised for lack of clarity in its meaning and dimensions and the question of how it should be implemented in practice.[2] Delgado and Åm (2018) state that the notions of societal concerns and public values remain vague. In practice, they note, RRI often comes down to "box-ticking", merely adding a paragraph to a research proposal and a social science and humanities (SSH) scholar to the team, usually last-minute. Another point of criticism concerns the meaning of the word "innovation". Von Schomberg and Blok (2021) argue that in EU innovation policy as well as standard RRI definitions, innovation is effectively equated with marketable technological products as in the definition quoted above ("the innovation process and its marketable products"). More systemic, social, cultural or economic innovations fall beyond the scope of this understanding. As long as innovation is equated with marketable technological products, the authority of the market as supreme decision-making authority will remain unquestioned, and success or failure will inevitably be measured in techno-economic terms. Objectives such as environmental sustainability, inclusion, and a liveable and equitable built environment do not compete on equal footing; they may be by-products but never conflicting objectives. The Socio-Technical Integration Research approach (STIR) (Fisher, 2007; Fisher & Schuurbiers, 2013; Fisher *et al.*, 2015) builds on the core idea of RRI that we can minimise unintended negative and maximise positive impacts of research and development and "nudge their trajectories in various ways toward responsible, desirable futures" (Stilgoe, 2013, p. 14) by adding an approach for midstream modulation of socio-technical research, strengthening reflection on the potential social implications of one's research during the research process itself. STIR has responded to many of the above criticisms, inter alia by providing a clear protocol which allows for in-depth interactions rather than mere box-ticking.

[2] For an overview see Burget *et al.*, 2017.

In this article, we use an adapted version of STIR to explore whether and to what extent it is possible to render the dominant techno-economic innovation paradigm (with its in-built market orientation) amenable to critical reflection and modulation in the midstream of research processes. To do so, we take our cues from the STIR approach, with its underlying interest in investigating the capacity of academic researchers to more reflexively attend to the social dimensions of their work and to align technoscientific and social considerations during the course of their research. We built on STIR as a rich reservoir of methodological knowledge, practical research experience and empirical findings concerning midstream modulation in interdisciplinary research collaboration. We followed the STIR method, as a semi-structured approach that in its deployment also requires considerable choices and interpretations, quite closely, however with a few adaptions to our case study. The aim of our research was to develop a deeper understanding of how researchers negotiate potential tensions between preset commitments to social and environmental values and values of efficiency, productivity and marketability and what the chances are to deliberately challenge the primacy of the latter through STIR interactions.

In this paper, we present results from two STIR case studies we conducted with researchers from two projects within a large German interdisciplinary research network on computational methods in architecture, engineering and construction. Through interactions with each participant in their everyday work over twelve weeks, we gained insights into the ways they dealt with the techno-economic paradigm. In the vast majority of cases, the STIR approach has already shown itself to be successful in terms of exploring the reasoning for research decisions as well as stimulating reflexive learning, value deliberations and practical adjustments (Schuurbiers, 2011; Lukovics & Fisher, 2017). We argue that in basic academic research, market requirements do not directly impact research but have an indirect impact through *anticipations* of market requirements. However, it is not *the* market that impacts research, or a single way of addressing assumed market requirements, but different ways of anticipating, addressing and approaching market requirements embedded in preset institutional contexts and research practices and different ways of negotiating them in relation to other values. Overall, we found that market requirements are only somewhat malleable and open to reflexive modulation.

In the remainder of this paper, we first introduce the STIR method and our adaptions of it for our specific research context. Subsequently, we present a condensed account of two STIR exercises as well as our findings with regard to anticipating and addressing market requirements and their respective embedding within particular project settings.

SOCIO-TECHNICAL INTEGRATION RESEARCH (STIR)

STIR stands in a long tradition of strategies for integrating societal concerns and considerations into technoscientific research and development, from ethical, legal and social implications (ELSI) and various strands of technology assessment (TA) to RRI (Fisher *et al.*, 2015; Kropp, 2021). Overall, as Job Timmermans concludes, these strategies have shown limited impact on actual developments in research and innovation. "RRI still is chiefly discussed conceptually in terms of frameworks and approaches rather than practically in terms of tools and knowledge transfer" (Timmermans, 2017, p. 20). Some strategies may have been successful in preventing the worst, but there is little indication that they have redirected research and innovation towards more sustainable and socially desirable ends. Integrating societal concerns has proven more difficult than expected, with obstacles and challenges arising from the dominance of the techno-economic innovation paradigm as well as rather formal and last-minute ways of involving SSH scholars and perspectives (Bogner *et al.*, 2015; Mayntz, 2015; Kuzma & Roberts, 2018; Manzeschke & Gransche, 2020; Stubbe, 2020; Strand *et al.*, 2021). STIR tackles these challenges to some degree: it provides a strategy for stimulating reflections on the social contexts of research through regular interactions between an SSH researcher and technoscientific researchers within a particular research setting over a certain period of time. These interactions, which usually take place over twelve weeks, allow for collaborative reflections on the social context and possible implications of the research and, potentially, opportunities for practical adjustments (Fisher *et al.*, 2006; Fisher & Schuurbiers, 2013). However, STIR aims not to radically reorient and change the course of research, but rather to incrementally take greater account of social, ethical and environmental assumptions in the research process wherever possible (Owen *et al.*, 2013).

Through inciting reflections on possible alternative research decisions and practices in the interaction between technology researchers and SSH researchers, STIR appears suited to motivate the contemplation of market requirements, their influence on the research process and the relationship between techno-economic imperatives and competing objectives of delivering socially desirable outcomes.

Thus, STIR is a form of what Konrad *et al.* (2017) term "constructive technology assessment" in that it...

> ...aims to mobilize insights on co-evolutionary dynamics of science, technology and society for anticipating and assessing technologies, rather than being predominantly concerned with assessing societal impacts of a quasi-given technology. In addition, it shifts the focus from policy advice to (soft) intervention in the ongoing construction and societal embedding of technologies. (Konrad *et al.*, 2017, p. 15)

Through documenting the engagement of technoscientific researchers with the socio-technical context of the work in real-time collaborative reflection about possible implications of their work, it is possible to analyse those co-evolutionary dynamics on a micro level and observe how the ongoing construction of technologies might change through such reflection (Schuurbiers, 2011).

STIR focuses on midstream modulation (MM) (Fisher *et al.*, 2006; Fisher *et al.*, 2015), seeking to distribute "responsibility *throughout* the innovation enterprise, locating it even at the level of scientific research practices" (Fisher & Rip, 2013, p. 175). MM denotes incremental processes of altering research and development practices and decision-making in response to social constraints as well as social values, considerations and influences more generally (Fisher *et al.*, 2006; Fisher & Schuurbiers, 2013; Owen *et al.*, 2013). Importantly, modulation occurs in any case; there is no unmodulated research or innovation process. The critical questions are whether modulation occurs consciously and whether it is oriented towards sustainability, equity and inclusion. STIR aims at MM as a way of enhancing the responsive capacity of researchers to consider the social and environmental contexts and implications of their work and, possibly, align research and innovation agendas more closely with public values and desirable futures (Fisher & Schuurbiers, 2013). It has been found to support two modes of reflexive learning: first-order and second-order. First-order reflexive learning focuses on reflecting and improving research decisions in relation to the framework of objectives, assumptions, background theories and values underlying the research, while second-order reflexive learning subjects the given normative, epistemological and institutional framework to critical inspection and reflection which then extend to the research culture and its methodological, epistemological, ontological and socio-ethical premises (Schuurbiers, 2011, p. 772).

We built on this approach but also interpreted it slightly differently in that every now and then, we explicitly raised questions regarding the social and environmental impacts of the research. We consider this appropriate since, in this case, the research network had already made a commitment to these goals, hence we did not bring them in from outside. By socio-technically integrating, jointly situating and critically interrogating the guiding research assumptions and practices in the sessions, our aims were to promote consideration of social, ethical and environmental aspects and impacts and to collectively explore possibilities for more than incremental adaptation to sustainability challenges.

STIR can result in first- and second-order reflexive learning through two modes of interaction, termed STIR 1.0 and STIR 2.0, between SSH researchers and those with science-technology backgrounds. STIR 1.0 has a more *reconstructive* character; it aims

to identify constraints and requirements, interests and expectations (such as the expectation to deliver marketable products) on the level of the research practices, institutions, or society at large, constituting the cultural background to the research process. STIR 1.0 probes and assesses the capacity of midstream agents to modulate their practices and research trajectories under given conditions (Fisher *et al.*, 2016). Therefore, STIR 1.0 can have an intervention-oriented and transformative effect on research practices and capacities. STIR 2.0 has a more explicitly *transformative* aspiration, attempting to deliberately modulate the research process by calling attention to social and ethical impacts and, with our adaptation of STIR, questioning given assumptions (Kropp, 2021). "As a research program, STIR 1.0 probes the conditions and capacities for broadening socio-technical integration while, as an intervention, STIR 2.0 attempts to exercise these capacities deliberatively" (Fisher *et al.*, 2015). Insofar as we deliberately invited the techno-scientific researchers to reflect upon the impacts of their research with regard to substantive values the research network had committed itself to, we indeed practiced a version of STIR 2.0. Yet, it was a version of STIR 2.0 in which the SSH researcher referred to values that had already been inscribed in the research framework; thus, the SSH researcher rather recalled these values and commitments in the interaction rather than newly introducing them.

STIR scholars have distinguished three types of MM, referring to different levels of reflecting and modifying research activities: *de facto*, reflexive and deliberate modulation (Flipse *et al.*, 2013). In *de facto modulation*, the SSH researcher recognises so-called decision modulators that shape the research process. Such modulators can encompass a variety of cognitive assumptions and social or material framework conditions that inform the research process, including guiding assumptions and expectations posed by the particular research settings or the institutional context. To which extent these may determine the research, and to which extent they can be modified, are empirical questions that cannot be answered in advance. In *reflexive modulation,* the participants become aware of the ways in which assumptions and expectations influence the actual research process, thereby making it possible to render them negotiable and modifiable. In *deliberate modulation*, researchers actively and deliberately integrate certain *de facto* modulations in their decision-making. (Fisher & Mahajan, 2006; Flipse *et al.*, 2013; Kropp, 2021). Following our adapted STIR approach, in *deliberate modulation,* we aimed for the researchers to not only deliberately integrate considerations about *de facto* modulations, but to question some and consider others addressed by the SSH researcher. Deliberate modulation is therefore of particular importance for responsible innovation. The three types of modulation also form conceptual devices for interpreting the data acquired through the STIR process (Flipse & van de Loo, 2018).

Based on our observations using the STIR protocol (see Fig. 2), we decided to expand the process by one step; we actively questioned and challenged guiding assumptions (such as techno-economic orientations) and asked participants to reflect on them. The participants were then invited to argue whether it was possible to more strongly orient their work towards social, ethical and environmental objectives beyond the dominant techno-economic paradigm. Our intention was to determine whether we could ascertain the participants' views on the potential of various modes of changing direction. The participants' accounts of the constraints reducing the possibilities of adjusting their work made it possible to understand the role of anticipated market requirements for research activities in the midstream.

STIR IN ARCHITECTURE AND CIVIL ENGINEERING

Our STIR exercises took place within an ongoing interdisciplinary research network in architecture and engineering. SSH researchers were included from the beginning and are contributing to the research.

Within this context, we collaborated with one researcher[3] from engineering (Researcher E) and one from architecture (Researcher A) in two separate STIR exercises. Both researchers assumed central roles within their respective projects.[4] During the two STIR periods, SSH Researcher I (for socio-technical integration) used so-called decision protocols (Fisher, 2007) and conducted twelve guideline-based semi-structured conversations with both researchers to jointly explore upcoming research decisions in terms of general objectives, options, guiding assumptions, possible alternatives and expected outcomes (see Figure 1). The decision protocol allows the participants to systematically address and reflect on their current research decisions against the background of the assumptions considered relevant and the possible courses of action. In this sense, the decision protocol serves as a guideline for these in-depth, problem-centred conversations. Through regular interaction and dialogue over several months, it was possible to track changes in the technology researcher's position and address recurring assumptions and themes in greater depth.

[3] STIR is not necessarily limited to interactions with individual researchers; however, due to the framework of the research context, our cooperation had to be closely coordinated in advance. In this study, we therefore limited STIR to individual interactions.

[4] For reasons of anonymity, we use the gender-neutral pronouns they/them/their when referring to the researchers.

Figure 1: STIR decision protocol as basis for semi-structured conversations

Opportunity Describe a problem, opportunity or decision you are facing.	Considerations What should you consider in responding to the opportunity?
What do you anticipate you will do, why, and who might care? Outcomes	What courses of action are available to you for responding? Alternatives

Source: Phelps & Fisher, 2011; authors' representation (Frost *et al.*, 2022). For the decision model underlying the decision protocol, see Fisher, 2007.

At the beginning and end of each STIR period, we conducted longer interviews on research goals, horizons and expected results. For the analysis, the decision protocol-based conversations were treated as semi-structured interviews. Researcher I also documented ethnographic observations during project meetings and experiments.[5] The research thus comprises 28 interviews as well as observations of seven meetings and experiments. All conversations were audio recorded and transcribed; the transcriptions, together with field notes and STIR decision protocols, were coded using MAXQDA. After open and selective coding, a case narrative was written and case information and different categories were cross-compared. The interviews with Researcher E were translated from German to English by Researcher I, and the translation was presented to Researcher E to confirm its accuracy.

We here present a condensed account of the STIR processes and their outcomes with regard to reflections on anticipated market requirements and their impact on the research process. Our account is based on a selection of significant passages from the STIR transcripts. Importantly, the fragments of socio-technical interactions presented here do not represent the entire range of themes and reflections within our STIR processes. Since the focus is on the role of anticipated market and industry requirements, we selected only those sequences that refer to this aspect.

Set goals: Sustainability, efficiency, design freedom

In the research projects studied, problem descriptions and objectives were not determined by the individual researchers but were provided by the research settings

specified in the research network. Moreover, both the interviews and Researcher I's observations during the course of the STIR exercises documented that the individual projects stood in the tradition of their respective research institutions and were shaped by previous research and innovation activities.

Department E's long-standing reputation for the development of lightweight concrete is reflected in the department's expertise, experiential knowledge, network of research partners, and technologies employed. It is of vital importance to the department to minimise resource consumption for building construction, as evidenced by publications, public lectures, and previous projects by department members. At the same time, the agenda of the overall research network emphasises the need to build more and faster to meet global needs. Material savings through lightweight construction are intended to (at least partially) reconcile these opposing challenges, to which Researcher E frequently referred. The development of lightweight building components reflects the dual objectives of reducing building mass and reducing CO_2 emissions in the production process. Moreover, prefabrication instead of on-site processing, as Researcher E explained, allows for components to be produced at higher speed, "because, of course, it's an industrial process and it thrives on you getting your products out quickly".

For Researcher E, developing resource-saving building components is a way to serve environmental, social and economic needs simultaneously:

> …because we have to build more and more quickly for the people who will soon all want to move out of their homes. […] And […], look here, this is a beam […] that has the same load-bearing capacity but a reduced weight. In the end, that is always a quality criterion for our own work to say we can do the same thing as is currently possible, but with less weight. That is always the primary guideline.

In Department A, as Researcher A pointed out, environmental concerns such as the excessive consumption of resources and the simultaneously growing demand for built-up space also play an important role. In addition, the aim is to increase design freedom and harness the full potential of computational technologies to create innovative and flexible building systems that would contribute to a more sustainable and liveable building culture. Researcher A consistently explained his research decisions in terms of creating new design options.

The issue of environmental sustainability also figured prominently in Researcher A's statements and is reflected in the research focus on timber buildings; timber is considered a renewable material which also stores CO_2. Researcher A explained that combining digital design and robotic manufacturing could increase design options as well as process efficiency and "make it high speed, high detail". Given these features, they expected resultant new building systems to be characterised by high flexibility and building longevity.

Department A draws on a history of material development, computer-based design and robotic manufacturing, as reflected in the technologies used and associated expertise. Researcher A mentioned numerous intensive, worldwide collaborations with other researchers in these fields; the department also maintains close contacts with the domestic construction industry and policy-makers, whose expectations were very familiar to the researcher.

As for the specifications set down by the overall research network, these refer explicitly to the global environmental crisis and the housing crisis in the context of global population growth. In light of these challenges, the network's overall objective is to promote digital technologies that advance a sustainable building culture and contribute to a high-quality, liveable and sustainable built environment. The overarching research framework includes the commitment to reduce CO_2 emissions and waste production as well as increase the productivity of building construction. Social and environmental goals thus stand side by side with techno-economic goals. Interdisciplinary research projects within the network have taken up this framework to define their specific research strategies and objectives, including resource-minimizing work with concrete and timber. These strategies and objectives were fixed and not to be implemented at the discretion of the individual researchers.

During the STIR exercises, Researchers A and E repeatedly expressed that they shared their departments' and the network's professed orientation towards social and environmental values. Alongside these values, and also in line with the network's overall objectives, process and system optimisation (particularly in terms of time efficiency) was another top priority for them. As these values can come into conflict with one another, they became a matter for reflexive modulation.

Disruptive change, incremental innovation, and the metrics of achievement

While the two projects started from similar problem descriptions and objectives, the researchers encountered different challenges to industry adoption of their proposed outcomes. Concrete is one of the most extensively used building materials in the world (Gagg, 2014), and the concrete construction industry is well established and stable. In contrast, timber buildings are on the rise but still make up only a small portion of buildings; the timber-building industry is comparatively small. Modular construction is popular in building with timber; it allows for relatively cost- and time-efficient mass production of building components but imposes considerable limits on form-finding and architectural design options.

In our case study, the researchers took the structures of the respective industries into account. Meeting market requirements was an indispensable prerequisite for them to implement a more sustainable construction method that

would gain ground beyond academia. The values of productivity, profitability and marketability therefore assumed first priority as they would determine whether the researchers' developments would eventually succeed. Both researchers anticipated and addressed these requirements but did so in differing ways, with one envisioning the diffusion of the prospective product within the given industry structure and the other proposing a sectoral change.

In architecture, the focus was on changing the building culture by disseminating new timber-building technologies that would disrupt the industry. As Researcher A explained:

> I would want it to be considered normal to build exclusively out of timber. […]. Whereas like in industry […] you would start designing something and then at some point somebody says, well, what's it made out of? Concrete or steel? And often timber [is] just fully left off the board. So, I think a positive change to the building culture would be for people to say to me, assuming it was made out of wood and then ask, "is it made out of wood or something else", you know, or at the very least for wood to be on this list, you know, like "Is it wood, concrete or steel?" Like three options rather than just two.

In engineering, the aim was to develop more sustainable but also techno-economically efficient solutions that would thus be attractive to the established industry. In this way, the outcomes of their research would make the industry more sustainable. For Researcher E, the focus was not on disruptive change but on incremental innovation:

> Therefore, our first goal is to make it [the building components] lighter. Because it is simply necessary, it is socially necessary. And after that, of course, it follows relatively quickly that people also have to apply it. And that means that it has to be easy to do, that I can perhaps even say that I don't have to completely change the industrial processes as they exist now. But I offer an addition. I offer a way to apply it differently for a better result.

In their research decisions, the two researchers thus adopted different strategies for dealing with the presumed market and industry requirements. Techno-economic criteria played a major role in the research decisions taken in Department E. Here, a key question was how to achieve material efficiency that would translate into time and cost efficiency. The technology under development had to have a competitive advantage over others in techno-economic terms; otherwise, it would assumedly have no chance of being adopted. Thus, the goal for Researcher E had to be "that we say we are better, we are lighter and cheaper. And faster in the end". The strategy was not to develop innovations for an industry that might have to adapt to future policy shifts. As Researcher E stated:

> […] the moment a carbon tax comes in, you've won with something like that. If you can then really say, we'll do the same, but [with] 50 percent, 60 percent less material.

In Research Project A, on the other hand, the strategy was to highlight the economic advantages of the novel research outcomes. These outcomes, the researcher explained, had to be comprehensible to industry actors and connect to existing knowledge, values and ideas about architecture and construction processes. Researcher A explained the idea by referring to reactions towards one of their earlier prototypes:

> It could have been a much more designed object, as this original version was. And that was sadly one of the feedbacks that we received when we showed this around, that it looks great, but it makes me think that it is a designed object, not that it is an example of a multi-storey building system.

Both researchers anticipated how the industries would respond to their work. In one respect, the researchers took a strategic stance towards the techno-economic imperative of increasing productivity, profitability and marketability; meeting this condition was seen as a means to the actual end, namely achieving a better, more sustainable building culture. In another respect, we see that, once the techno-economic imperative is accepted, it tends to outweigh all other ends and values and to define the very standards for measuring success and achievement. Researcher A, perhaps unwittingly, illustrated this mechanism:

> And if we can even target five percent of buildings, it's a 13 trillion euros per year market. OK, and even five percent of 13 trillion is great. So [I'm] cool with that number.

Making money was certainly not the purpose of their research; advancing sustainability and freedom of design was. Still, achievement and success were gauged in terms of money. Why is this? Perhaps we encounter a more fundamental problem here that occurs when qualitative goods, such as environmental sustainability or freedom of design, compete with quantifiable ones, such as market share and market volume. Before relative weights can be assigned to qualitative and quantitative goods, a common standard must render them comparable. A common solution is to translate quality into quantity – social, cultural or environmental values into economic ones. In this case, the researcher sought to express the importance and desirability of non-quantifiable values by translating them into economic ones. Such economic valuation of environmental principles is not an individual process derived from external constraints but a long-established routine relying on social agreements based on valuations (Prior, 1998; Asdal, 2015). Yet, strictly speaking, this translation is logically impossible; quality *is not* quantity, and to quantify qualitative goods actually means to negate the difference. Qualitative goods are then taken into account only if they are valuated or "co-modified" (Asdal, 2015, p. 169-170) in relation to quantifiable ones.

For this reason, we would argue, social, cultural and environmental values literally cannot compete with economic ones. In the tradition of seeking optimal resource allocation, however, only the quantified is governable; accordingly, this translation is a routine technique in innovation processes.

On the whole, both researchers taking part in our STIR exercises took for granted that market and industry requirements could not be suspended or circumvented. In STIR terminology, these requirements were considered to be beyond the scope of MM and thus beyond the reach of collaborative modifications — a finding that points to the need for RRI activities on striving for an upstream level of innovation policy by, for example, providing effective incentives (Gurzawska *et al.*, 2017; Manzeschke & Gransche, 2020).

In the following section, we discuss the extent to which STIR has proven suitable for rendering conflicting research objectives – in particular techno-economic *versus* social, cultural or environmental objectives – amenable to reflexive modulation. We present two instances in which STIR 2.0 was practised to challenge the primacy of techno-economic objectives in research.

ADDRESSING CONFLICTING OBJECTIVES – "There is little we can do to affect their economics"

Let us consider conflicting research goals as they occurred in the course of STIR interactions in Project A. Researchers I and A were discussing the possibility of restricting design options to those that meet the so-called Goldilocks density – dense enough but not too high – to save land and encourage the liveability and affordability of cities. While developing technologies to meet the Goldilocks density might be a socially desirable goal, Researcher A explained, it would conflict with the goals of increasing freedom of design, demonstrating economic benefits, and presenting design options for timber construction – that is, goals that had been fixed in the overall framework for the research project:

> So, if we had talked about restrictions like the seven-storey thing, it would have reduced our potential for impact. […] There exist opinions that this Goldilocks density is correct for urban life — I don't know if I hundred percent agree with them — but I think they're quite nice. But we would not want to restrict anything we are designing or anything we are building to that. I think that the more types of buildings and the more heights and sizes and shapes of buildings that are possible, the better. It supports the thesis which is expanding what is possible within timber building construction.

Here, the goal of expanding design options for timber buildings and demonstrating them was given priority over promoting buildings that promise socially desirable urban density. Decisions regarding building height and density would be left to future

construction actors such as clients or planners; the researcher would not need to determine these parameters in their own developments. In this case, the conflict between different research objectives – expanding design options and demonstrating the varietal range of timber buildings on one hand and contributing to a more liveable and sustainable spatial order on the other – was managed by dividing the responsibility between the innovator (for generating design options) and future construction actors (for deciding which ones to realise). In short, we could say that reorienting the research towards more desirable social outcomes was discarded in favour of relying on established but as yet unsuccessful downstream modulation via regulations and market mechanisms: "*De facto* policies of hoping for the best and letting the future take care of itself" (Stilgoe, 2013, p. xv) are widespread, STIR interactions notwithstanding.

Another conflict emerged between the objective of developing more durable and therefore sustainable buildings on one hand and optimizing process efficiency on the other. For Researcher A, developing efficient design and production methods as well as durable, hence sustainable buildings were a major objective around which all research decisions were oriented. Process efficiency and building longevity, however, can collide when increased process efficiency makes it profitable for investors to demolish existing buildings and quickly build new ones in large numbers. This would be a case of an unintended rebound effect: an individual new building might be environmentally sound, but economic incentives can lead to increased construction activities that outweigh the benefits of sustainable building.

Researcher A was aware of this potential collision of goals but saw no way to address it in their work as "there is little we can do to affect their economics". From their point of view, it was beyond the scope of their research, and consequently the need to assign priority to one of these conflicting objectives did not arise. While this merely eschewed the conflict, Researcher A resorted to another solution, concluding that market mechanisms may drive the premature demolition of buildings and that, given these mechanisms, increased process efficiency could even reinforce this tendency:

> I would say that many of the buildings that are considered thrown away or that are built […] for 20-year lifespans and destroyed after five are buildings of lower quality. And by quality, I mean not only that their materials are cheap, but their design is simple.

Another outcome, though – increased design quality – would outweigh that adverse effect. In this case, Researcher A chose to reduce the potential for unintended effects of innovation and resolved the tension between the two objectives in a way that was compatible with the general framework settings outlined above. These stipulate that

increasing building longevity and process efficiency and design options are compatible. Within this framework, the techno-economic innovation paradigm remains unquestioned; conflicts between the techno-economic goals of increasing time and cost efficiency and the social, cultural and environmental goals of a sustainable, liveable built environment cannot be addressed as long as achieving the former is taken as a precondition for the latter. Consequently, the researcher's scope for responsible decision-making is seen as confined by the laws of the market; they are aware of these constraints but do not see a way or need to modulate them. In this case, the adapted STIR approach with its transformative aspiration did not stimulate critical debate on the primacy of the techno-economic paradigm, its influence on research and possible undesired effects of that research.

Is promoting social justice feasible?

The question of undesired side effects also arose in Research Project E. During STIR interviews, Researchers I and E discussed the question of whether new methods to reduce material consumption in construction might reinforce the trend to build bigger and more. In a global context, Researcher E argued, this raises questions of social justice:

> Because the question is whether we are allowed to emit more CO_2 now, for example, just because we can, because we have the space, *de facto*, we have the space for it, we can emit more CO_2 than in New Delhi when new buildings are being constructed there, because it is simply less space per capita available. It's a question of justice that comes up again and again.

A responsible decision, the researcher argued, would be to refrain from building in many places in which one could. Yet, again, they saw no way of incorporating these considerations into their everyday work; the issue seemed to be beyond the scope of the micro-decisions made in research practices:

> The problem is not that I don't like to acknowledge that it would make moral sense for us to take less so that others can get more first, until a point where we say, okay, now we're kind of on the same level. I would regard the problem as one of feasibility.

The researcher was aware that their research might have the unintended effect of further fuelling construction activities and, in turn, land use and CO_2 emissions, thereby possibly exacerbating existing global justice problems. Still, it did not seem feasible for affluent countries to reduce construction activities and CO_2 emissions for the benefit of those in other countries. In any case, the researcher did not consider it necessary to deliberate the question in their actual research; they kept instead to the more obvious and realistic option of making construction more material-, cost- and time-efficient. While it is certainly debatable to what extent such a highly complex,

global problem can be taken into account in academic research, it was striking that there was no further attempt to discuss possible ways out of this dilemma. Suggestions by Researcher I to think, for example, about ways of developing a building system for smaller layouts were not taken further.

These interactions illustrate that STIR can indeed inspire reflections on responsible research; the participants considered the social and environmental implications and the side effects of their work. Yet these considerations remained somewhat abstract in terms of effects attributable to market imperatives such as increasing profits by building more in less time. They did not know how to integrate these aspects into their everyday work as they seemed unmanageable, far removed from their own sphere of influence and beyond their perceived responsibilities. When social and environmental objectives such as building longevity, flexibility of use, and reducing CO_2 emissions conflicted with techno-economic goals, the latter always prevailed, less due to conscious decision-making and priority-setting than to the underlying assumption that there was no way around market mechanisms and that only those research outcomes that meet private-sector economic requirements could have any impact. In other words, research outcomes could translate into successful innovations only if they were to provide a demonstrable techno-economic benefit proven in the market. Through our adaption of STIR 2.0, which challenged the assumption that downstream dynamics could not be influenced by midstream activities, it became clear that the participants found it both unrealistic and inadvisable to neglect market requirements in favour of building durable structures and promoting environmental and social justice.

QUESTIONING ASSUMPTIONS IN DECISION-MAKING

Ideally, STIR opens up spaces for second-order reflexive learning processes, that is, reflections on the possible social implications of one's research, even if these might challenge the underlying assumptions and expectations of the research settings or the societal context. In the following sections, we analyse instances of second-order reflexive learning in the context of STIR 2.0, which, as described above, we sought to achieve by critically questioning the primacy of the techno-economic innovation paradigm in reaction to the explicit sustainability commitments made by the overall research network. In the first example, Researcher I questioned the concept of co-design, which was key to the work of Researcher A. In the second example, the concept of democratic digitalisation discussed during STIR sessions modulated the assumptions underlying Researcher E's decision-making process. While the first example points to possible barriers to deliberate modulation in terms of responsible

innovation, the second shows how the integration of socio-ethical concepts in the decision-making process can succeed.

STIRring the concept of co-design

Researcher A's work on a multi-agent system for computational design of multi-storey timber buildings was strongly influenced by the research network's understanding of co-design. Simply put, co-design denotes an approach for integrating design and construction processes through computer-based feedback loops. It entails multidisciplinary collaboration among construction professionals from various fields such as architecture, structural design, building physics and lifecycle assessment.

Integrating the needs or expectations of further stakeholders was not a constitutive element of Researcher A's co-design concept. The focus was on integrating technical and environmental requirements and related professional expertise, not because the requirements of residents or stakeholders were deemed irrelevant but because they seemed incompatible with the computer-based, numerical approach that had been chosen. As Researcher A explained:

> So, if somebody else would want to use this similar approach, but then maybe also integrate [....] retail consultant knowledge or something like this, then, you're right, they would not be able to use this tool. If it is possible that they would, why would they even want to? So, I think this brings up a harder question about what to try to automate, because at least from my understanding and my experience within architecture for the last decade, the key data-driven slash numerical players in every project are some amount of structural designer, the architect and the material use [referring to lifecycle assessment]. So that is why I do think that the reason we've included building physics into this as well is because it is a highly numerical, data-based field.

Non-numerical matters of concern escaped the technologies already developed and could not be integrated into the new approach. From the perspective of science and technology studies (STS) (Jasanoff, 2004), we can see this as a case of co-production in the sense that the technological approach co-defined the social and cultural values at stake – technical and environmental quality, not stakeholder or community participation – and co-shaped the social practices of planning, constructing and inhabiting buildings. Schikowitz (2020, p. 222) points out that the production of societally relevant knowledge to which researchers aspire must make research "do-able by aligning diverging commitments, concerns, requirements and practices", especially in situations of conflicting objectives. In this case, the alignment seems to be accomplished through the reliance of researchers on quantitative and numerical approaches.

On one occasion, a discussion arose between the social scientist and the architectural researcher on whether to expand the co-design method to include

community or stakeholder perspectives through defined interfaces. Researcher A argued that these matters were important but should be left to the planning architect:

> [This] is a thing that the architect is supposed to do, to kind of handle how the product or how the building will affect the community. And I think that this continues to be an architect-specific task and not a community-based task. So, the community will express its views or desires, and it is up to the architect to distil them and implement them within this co-design. Within my understanding of design, the vox populi will still be heard, but it will be heard through the filter of the architect. And that is how it will affect the design. I don't think it affects the co-design directly.

The researcher here evokes the notion of the architect as the mastermind of the construction process who integrates all requirements. From their perspective, freedom means first and foremost freedom of architectural design. In terms of co-production, one could say that a technological approach oriented towards aspects and activities that can be represented on a numerical basis was perceived as compatible with the objectives of expanding architectural design options and increasing system efficiency. The linkage between these elements proved quite stable throughout the STIR process and was not challenged by deliberate reflection. Integrating aspects of non-numerically representable quality through participatory design options, under these premises, appeared unfeasible. Using Schikowitz's concept of the production of societally relevant knowledge, we can understand this episode as an effort by the researcher to make the different commitments, concerns, demands and practices do-able by resorting to the quantitative methods and goals available to them (Schikowitz, 2020).

"Good" digitalisation: "You don't have to imitate everything"

In Project E, STIR interactions prompted discussions on the social dimensions of the research on a rather fundamental level. Arguably, such discussions can influence research decisions and in this sense lead to deliberate modulation. One such instance referred to a conversation on "good" digitalisation and the framing of a research decision by Researcher E in terms of social implications.

As mentioned above, Project E was working on a cyber-physical system for producing material-saving building components. During the STIR process, a project decision was taken to employ a modular system in order to enable separate instalment of individual pieces of equipment such as a laser scanner or an automated extruder. This would allow users to automatize some components of the production process while keeping the manual nature of other parts in place for the time being. Researcher E explained that a modular system would lower the cost of investment and give users more flexibility:

> The question is: Do you have to buy the entire system? […] Because that's the problem, then you have another juggernaut, and either you have the thing or you don't. But if I say that the whole system can also be 3, 4, 5, 6 individual modules, which I then can possibly link with each other, then the hurdle of getting a single module and achieving an improvement is lower, and probably you can then also optimise each module individually.

The topic of user flexibility and the question of when and what to automate harked back to a previous remark by the researcher and a STIR discussion on "good" digitalisation:

> That's why I think it's good that we have this discussion [about 'good' digitalisation] in our [research network], that we can also ask ourselves, do we have to go in this direction now and in which direction do you start and where does digitalisation really bring added value?

Researcher E added that "good" digitalisation must be "[distinguished] from a capitalist or capitalistically-shaped digitalisation [and] from a dictatorial or dictatorship-shaped digitalisation", referring to digital surveillance technologies, which the researcher considered problematic from a democratic point of view. In the context of this conversation, Researchers I and E discussed what would constitute "good" digitalisation, when automation would make sense from a more-than-techno-economic point of view, what good work in cyber-physical production systems could mean and what ethical and social problems could arise from uncontrolled digitalisation. They agreed that not everything should be digitalised or automated:

> Researcher E: I think you can do it right and you can do it wrong. […] But in my opinion you don't have to copy these things [that are done in China or the US], but you can also say: How should a democracy actually look like, how should digitalisation […] actually look like in an open and free democracy? […] But you have to ask yourself, what does digitalisation look like in our country? So how would it work in our culture? Researcher I: Do we want it, in what form do we want it? For what purpose? Researcher E: Yes, exactly, like that, and then you don't have to imitate everything, you don't have to [digitalise] everything.

We observe here a kind of reflexive learning for socially robust innovations and, as such, a case of deliberate modulation emerging from conversations about good, democratically desirable digitalisation that led to the decision for a modular system and were reflected in the accompanying rationale. At the same time, the rationale for the modular approach conflated the question of what made sense for society with the question of what made sense for businesses; democratic, socially desirable digitalisation was represented in economic terms, and again we see that social considerations can be integrated when they appear not opposed to, but compatible with market requirements.

CONCLUSION

Within the techno-economic innovation paradigm, socially and environmentally responsible outcomes are assumed to follow from techno-economic innovation. In this paper we have explored the chances and limitations of socio-technical integration when it comes to challenging the primacy of that paradigm in academic research.

In particular, our question was what the chances and limitations are of provoking reflections on possible conflicts between values of efficiency and productivity and social and environmental values in the research process, thus challenging the primacy of the techno-economic innovation paradigm. To do so, we applied a slightly modified version of STIR, STIR 2.0 as we put it, by critically questioning assumptions and objectives that conflict with social and environmental research goals. The STIR processes we conducted within two projects have shown the techno-economic innovation paradigm to be rather resilient towards deliberate modulation. Questioning such fundamental orientations, we conclude, does not suffice to mitigate the influence of market imperatives in the research process; midstream modulation is not sufficient to put other concerns on the agenda vis-à-vis these external expectations which are deeply rooted in and incentivised by scientific institutions. Obviously, a truly multi-level systemic change also requires upstream and downstream modulation and therefore requires broader governance of knowledge production involving governmental bodies, industrial and civil society actors to address market deficits (von Schomberg & Hankins, 2019, p. 2).

While these findings will not come as a complete surprise to Socio-Technical Integration researchers, this article has shown *how* midstream actors deal with the tension between market imperatives and techno-economic values on the one hand and social and environmental values and commitments on the other. In particular, we observed certain patterns of how researchers sought to negotiate these tensions.

We can recognise one underlying assumption at work and two ways of dealing with situations of tension in which the techno-economic paradigm is challenged. The tacit assumption underlying the overall research framework as well as the individual projects was that marketability of prospective outcomes was not one objective amongst others but the precondition for all others. According to this assumption, marketability is not everything, but without it, everything would be nothing, given that sustainable, socially and aesthetically attractive products can only make a difference if adopted by the market. Therefore, societally desirable goals and techno-economic ones did not compete on equal footing; in case of conflict, the prioritisation of the former was always pre-determined. This assumption, which characterises the techno-economic innovation paradigm in general (Callon, 2002), could not be fundamentally challenged through our STIR exercises.

The participants did reflect upon the social implications of their work, including possibly undesirable effects arising from market mechanisms (incentive to demolition, aggravating global injustice). In some instances, modifications of specific research concepts (co-design) or strategies (systemic or modular automation) in order to integrate social concerns were considered. In these situations, the possibility of conflicting goals emerged, yet was resolved through two recurring argumentative patterns which we call the *lack-of-agency* and the *reconciliation-after-all* patterns. By "lack of agency" we mean that a particular course of action was considered unfeasible on technical and/or economic grounds and deemed beyond the sphere of influence of the individual researcher (e.g. for or against Goldilocks density, building demolition, increased construction activity). Another way of managing such conflicts was to point out how the research would contribute to reconciling them in the future "after all" (e.g. building quality superior to that of those demolished). Deliberate modulation in favour of societal concerns thus occurred, but only when it aligned with market requirements. As a result, the chances of critically challenging the above-mentioned tacit assumption and initiate substantive changes through social considerations proved to be limited. Nevertheless, our STIR 2.0 exercises have led to a better understanding of how anticipated downstream in form of market requirements is stabilised in research practices. We were able to observe how co-production of technological approaches co-defined the social and cultural values at stake and co-shaped the social practices of research. At the same time, the relevance of making technology research do-able by aligning diverging commitments, concerns, requirements and practices became quite clear.

These findings point to a lack of alternatives to market-driven diffusion of research outcomes, making it difficult for researchers to visualise success and achievement independent of market requirements. If the aspiration of research is change through adaptation of research results in a particular field, researchers appear to perceive the anticipated downstream as being of enormous importance. Every development, no matter how socially desirable and ecologically sustainable, must then prove itself capable of competing according to the logic of the field. This indicates that an upstream agenda alone, even one with an explicit normative social and environmental commitment as in these cases, can only succeed alongside critical consideration of the downstream. Only by simultaneously considering and shaping all moments of the innovation process can more socially responsible and sustainable development pathways be conceived. Policy-derived, subsidised niches for sustainable innovations, such as those Germany has created for renewable energies, are promising, but they are as yet scarce in the construction sector. Thus, in order to reorient research in digital architecture and construction away from techno-economic imperatives, socio-technical integration would have to confront them at all levels:

midstream, upstream and downstream (Fisher & Schuurbiers, 2013; cf. Yaghmaei & van de Poel, 2021). Creating niches as incubation rooms for radical novelties, locations for learning processes and space to build supportive social networks (Geels, 2002, p. 1261) may counteract the recurring "lack-of-agency" and "reconciliation-after-all" argumentative patterns in research practice. Even if niches cannot escape techno-economic imperatives, they are to a certain extent protected spaces for experimentation that have a more open character of configurability and do not require an immediate connection to given regimes such as market requirements in order to innovate.

Let us again emphasise that the problem is not with technological efficiency as such; difficulties arise, however, when it is defined and measured solely in economic terms, with social and environmental values only considered to the extent that they can be translated into the former. STIR and our approach of STIR 2.0 can evidently create awareness of this structural problem but not solve it. Or, conversely, socio-technical integration in research cannot solve the problem, but it can create awareness of it. The space for midstream modulation, in the cases we studied, was shaped by gateways that were opened further upstream, and the anticipation of market and industry reactions to be encountered further downstream. In accordance with the STIR literature, this article shows that greater efforts are needed beyond midstream constellations in order to bring about a departure from techno-economic imperatives in technoscientific research. Moreover, however, by pointing out the "lack-of-agency" and the "reconciliation-after-all" patterns of argumentation, the article could shed some light on *how* the techno-economic innovation paradigm is able to become resilient towards critical questioning, thus stabilizing upstream and downstream imperatives within midstream research practices.

REFERENCES

Asdal, K. (2015). Enacting values from the sea. On innovation devices, value practices, and the co-modification of markets and bodies in aquaculture. In I. Dussauge, C.-F. Helgesson & F. Lee (Eds.), *Value Practices in the Life Sciences and Medicine* (pp. 168-185). University Press Scholarship Online.

Braun, K., & Kropp, C. (2021). Schöne neue Bauwelt? Versprechen, Visionen und Wege des digitalen Planens und Bauens. In K. Braun & C. Kropp (Eds.), *In digitaler Gesellschaft. Neukonfigurationen zwischen Robotern, Algorithmen und Usern* (pp.135-165). transcript.

Bogner, A., Decker, M., & Sotoudeh, M. (2015). Technikfolgenabschätzung und Responsible Innovation. In A. Bogner, M. Decker, & M. Sotoudeh (Eds.), *Responsible Innovation. Neue Impulse für die Technikfolgenabschätzung?* (pp. 11-28). Nomos.

Burget, M., Bardone, E., & Pedaste, M. (2017). Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and*

Engineering Ethics, 23(1), pp. 1-19. https://link.springer.com/article/10.1007/s11948-016-9782-1

Callon, M. (2002). From Science as an Economic Activity to Socioeconomics of Scientific Research: The Dynamics of Emergent and Consolidated Techno-economic Networks. In P. Morowski & E. M. Sent (Eds.), *Science Bought and Sold. Essays in the Economics of Science* (pp. 277-317). University of Chicago Press.

Coroamă, V. C., & Mattern, F. (2019). Digital Rebound – Why Digitalization Will Not Redeem Us Our Environmental Sins. *Proceedings of the 6th international conference on ICT for Sustainability, ICT4S* 2019 (2382).

Delgado, A., & Åm, H. (2018). Experiments in interdisciplinarity: Responsible research and innovation and the public good. *PLoS biology,* 16(3), e2003921. https://doi: 10.1371/journal.pbio.2003921

Fisher, E., Mahajan, R. L., & Mitcham, C. (2006). Midstream modulation of technology: governance from within. *Bulletin of Science, Technology & Society,* 26(6), 485-496. https://doi: 10.1177/0270467606295402

Fisher, E. (2007). Ethnographic Invention: Probing the Capacity of Laboratory Decisions. *Nanoethics* 1, 155-165. https://doi: 10.1007/s11569-007-0016-5

Fisher, E., & Schuurbiers, D. (2013). Socio-technical integration research: Collaborative inquiry at the midstream of research and development. In N. Doorn, D. Schuurbiers, I. van de Poel & M. E. Gorman (Eds.), *Early engagement and new technologies: Opening up the laboratory. Philosophy of Engineering and Technology* (vol 16, pp. 97-110). Springer. https://doi.org/10.1007/978-94-007-7844-3_5

Fisher, E., O'Rourke, M., Evans, R., Kennedy, E. B., Gorman, M. E., & Seager, T. P. (2015). Mapping the integrative field: Taking stock of socio-technical collaborations. *Journal of Responsible Innovation,* 2(1), 39-61. https://doi: 10.1080/23299460.2014.1001671

Fisher, E., Konrad, K.E., Boenik, M., Schulze Greiving-Stimberg, V.C., Walhout, B. (2016). Building an Agenda for Socio-Technical Integration Approaches. In D. M. Bowman, A. Dijkstra, C. Fautz, J. S. Guivant, K. Konrad, K., H. van Lente & S. Woll (Eds.), *Responsibility and Emerging Technologies: Experiences, Education and Beyond* (pp. 43-56). Amsterdam.

Flipse, S. M., Van Der Sanden, M. C., & Osseweijer, P. (2013). Midstream modulation in biotechnology industry: Redefining what is 'part of the job' of researchers in industry. *Science and Engineering Ethics,* 19(3), 1141-1164. https://doi: 10.1007/s11948-012-9411-6

Flipse, S. M., & Van De Loo, C. J. (2018). Responsible innovation during front-end development: increasing intervention capacities for enhancing project management reflections on complexity. *Journal of Responsible Innovation,* 5(2), 225-240. https://doi: 10.1080/23299460.2018.1465168

Gagg, C. R. (2014). Cement and concrete as an engineering material: An historic appraisal and case study analysis. *Engineering Failure Analysis,* 40(5), 114-140. https://doi: 10.1016/j.engfailanal.2014.02.004

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy,* 31(8-9), 1257-1274. https://doi.org/10.1016/S0048-7333(02)00062-8

Gurzawska, A., Mäkinen, M., & Brey, P. (2017). Implementation of Responsible Research and Innovation (RRI) Practices in Industry: Providing the Right Incentives. *Sustainability,* 9(10), 1759. https://www.mdpi.com/2071-1050/9/10/1759

Jasanoff, S. (2004). The idiom of co-production. In S. Jasanoff (Ed.), *States of Knowledge. The co-production of science and social order* (pp. 1-13). Routledge.

Joly, P.-B., & Rip, A. (2012). Innovationsregime und die Potentiale kollektiven Experimentierens. In G. Beck & C. Kropp (Eds.), *Gesellschaft innovativ. Wer sind die Akteure?* (pp. 217-233). VS Verlag für Sozialwissenschaften.

Konrad, K., Rip, A. & Greiving-Stimberg, V. S. (2017). Constructive Technology Assessment–STS for and with technology actors. *EASST review,* 36(3), 15-19. https://easst.net/wp-content/uploads/2017/11/review_2017_11.pdf

Kropp, C. (2021). Embedded Humanism: Chancen und Risiken von STIR für eine transformative TA. In R. Lindner, M. Decker, E. Ehrensperger, N. B. Heyen, S. Lingner, C. Scherz & M. Sotoudeh (Eds.), *Gesellschaftliche Transformation: Gegenstand oder Aufgabe der Technikfolgenabschätzung* (22nd ed., pp. 119-131). Nomos.

Kuzma, J., & Roberts, P. (2018). Cataloguing the barriers facing RRI in innovation pathways: a response to the dilemma of societal alignment. *Journal of Responsible Innovation,* 5(3), 338-346. https://doi.org/10.1080/23299460.2018.1511329

Lange, S., Pohl, J. & Santarius, T. (2020). Digitalization and energy consumption. Does ICT reduce energy demand? *Ecological Economics*, 176, 106760. https://doi.org/10.1016/j.ecolecon.2020.106760

Lee, G., & Borrmann, A. (2020). BIM policy and management. *Construction Management and Economics* 38 (5), pp. 413-419. https://doi.org/10.1080/01446193.2020.1726979

Lukovics, M., & Fisher, E. (2017). Socio-technical integration research in an Eastern European setting: Distinct features, challenges and opportunities. *Society and Economy*, 39(4), 501-528. https://doi.org/10.1556/204.2017.004

Manzeschke, A., & Gransche, B. (2020). Aufs Ganze gesehen. In B. Gransche, & A. Manzeschke (Ed.), *Das geteilte Ganze* (pp. 235-347). Springer Fachmedien.

Mayntz, R. (2015). Technikfolgenabschätzung – Herausforderungen und Grenzen. In A. Bogner, M. Decker & M. Sotoudeh (Eds.), *Responsible Innovation. Neue Impulse für die Technikfolgenabschätzung?* (1st ed., pp. 29-46). Nomos Edition Sigma (Gesellschaft, Technik, Umwelt, Neue Folge). https://doi.org/10.5771/9783845272825-29

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: Public Affairs.

OECD (2019). *Global Material Resources Outlook to 2060: Economic Drivers and Environmental Consequences.* OECD Publishing. https://doi.org/10.1787/9789264307452-en

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy,* 39(6), 751-760. https://doi.org/10.1093/scipol/scs093

Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A framework for responsible innovation. In M. Heintz, R. Owen, & J. R. Bessant (Eds.), *Responsible Innovation* (1st ed., pp. 27-50). John Wiley & Sons.

Phelps, R., & Fisher, E. (2011). Legislating the Laboratory? Promotion and Precaution in a Nanomaterials Company. In S. J. Hurst (Ed.), *Biomedical Nanotechnology. Methods and Protocols* (pp. 339-358). Human Press.

Prior, M. (1998). Economic Valuation and Environmental Values. *Environmental Values,* 7(4), 423-441.

Ribeirinho, M. J., Mischke, J., Strube, G., Sjödin, E., Blanco, J. L., Palter, R., Biörck, J., Rockhill, D., & Andersson, T. (2020). *The next normal in construction. How disruption is reshaping the world's largest ecosystem.* McKinsey & Company.

Roland Berger (2016). *Digitization in the construction industry. Building Europe's road to "Construction 4.0"*. Roland Berger GmbH.

Schikowitz, A. (2020). Creating relevant knowledge in transdisciplinary research projects – Coping with inherent tensions. *Journal of Responsible Innovation*, 7(2), 217-237. https://doi: 10.1080/23299460.2019.1653154

Schuurbiers, D. (2011). What happens in the lab: Applying midstream modulation to enhance critical reflection in the laboratory. *Science and Engineering Ethics,* 17(4), 769-788. https://doi: 10.1007/s11948-011-9317-8

Stilgoe, J. (2013). Foreword: Why Responsible Innovation? In M. Heintz, R. Owen, & J. R. Bessant (Eds.), *Responsible Innovation. Managing the Responsible Emergence of Science and Innovation in Society* (pp. xi-xvi). Wiley.

Stubbe, J. (2020). Sechs Thesen für gelingende Integrierte Forschung. In B. Gransche & A. Manzeschke (Eds.), *Das geteilte Ganze* (pp. 197-211). Springer Fachmedien.

Timmermans, J. (2017). Mapping the RRI Landscape: An Overview of Organisations, Projects, Persons, Areas and Topics. In L. Asveld, R. van Dam-Mieras, T. Swierstra, S. Lavrijssen, K. Linse, & J. van den Hoven (Eds.), *Responsible Innovation 3* (pp. 21-47). Springer International Publishing.

UN (2019). *World Population Prospects 2019: Highlights*. UN Department of Economic and Social Affairs, Population Division. ST/ESA/SER.A/423.

UN (2020). *Policy Brief: COVID-19 in an Urban World*. United Nations.

UNEP (2020). *Global Status Report for Building and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector.* United Nations Environment Programme.

Von Schomberg, R. (2008). Prospects for Technology Assessment in a framework of responsible research and innovation. In M. Dusseldorp & R. Beecroft (Eds.), *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden,* (pp. 39-61). VS Verlag.

Von Schomberg, L., & Blok, V. (2021). Technology in the Age of Innovation: Responsible Innovation as a New Subdomain Within the Philosophy of Technology. *Philosophy & Technology*, 34, 309-323. https://doi:10.1007/s13347-019-00386-3

Von Schomberg, R., & Hankins, J. (2019). Introduction to the International Handbook on Responsible Innovation. In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation. A Global Resource* (pp. 1-11). Edward Elgar Publishing.

Yaghmaei, E., & Van De Poel, I. (2021). *Assessment of Responsible Innovation: Methods and Practices*. Routledge.

Zhang, L., Balangé, L., Braun, K., Di Bari, R., Horn, R., Hos, D., Kropp, C., Leeistner, P., & Schwieger, V. (2020). Quality as Driver for Sustainable Construction – Holistic Quality Model and Assessment. *Sustainability,* 12(19), 7847. https://doi.org/10.3390/su12197847

A Normative Understanding of Innovation

Kevin H. Michels*

* College of New Jersey

ABSTRACT

Commentators have bemoaned the absence of a clear conceptual understanding of innovation both generally and within responsible innovation (RI). Much of our thinking about innovation is fragmented into separate categories such as "business," "social" or "technological" innovation with no clear understanding of the term these adjectives modify. In addition, RI discussions focus overwhelmingly on technological advances delivered through the marketplace, which are only a portion of the innovation story. Clearly, we need to develop a stronger account of the concept of innovation. What criteria must be satisfied for a contribution to the world to qualify as an innovation or, more simply, what is an innovation? This article will contend that innovation is inescapably normative, and that we can construct an understanding of innovation by elaborating on its normative elements and their implications. Innovation, I will propose, is *ethical change that delivers substantial applied value to beneficiaries of a domain*. After developing this account, I will show how it can reframe our understanding of innovation's relationship with technology and the marketplace, the innovator's understanding of technology, who gets to innovate, and why the various categories of innovation may be more diverting than helpful. I will also reflect on how the account of innovation offered here can refine our understanding of RI.[1]

Keywords: Innovation; Normative; Values; Ethics; Technology; Responsible Innovation; Definition.

Proposal Submitted 30 August 2020. Article Received 9 April 2021. Reviews Delivered 20 May 2021. Revised 27 July 2021. Accepted 20 August 2021. Available online 28. February 2022.

[1] Acknowledgement: Thanks to Kathryn A. Hockenjos and to NOvation reviewers Andrew Maynard and Gérald Gaglio for their thoughtful comments.

INTRODUCTION – CHALLENGE AND OVERVIEW

Commentators have bemoaned the lack of a clear conceptual understanding of innovation both generally and within the framework of responsible innovation ("RI") (von Schomberg & Blok, 2018). As Blok (2021) notes, innovation is "nowadays self-evidently understood as the commercialization of technological inventions". By this understanding, how are we to understand the contributions that are: technological but not commercial, such as Merck's donation of a drug to cure river blindness (Merck, 2021); commercial but not technological, such as business model innovation (Johnson, 2018); or neither, such hospice care (Parkes, 2008)?

Gaglio, Godin & Pfotenhauer (2019) describe (without endorsing) the profusion of innovation categories as "X-innovation" – social, technological, industrial, organizational, open innovation, and so on. But how are we to understand *innovation* – the term that the X terms modify? The lack of a more general understanding of innovation may help explain the proliferation of these domain- or method-dependent understandings of the term. Is there something that all innovations – technological, social, commercial, or otherwise – have in common? A *unified* understanding of innovation may show these X categories to be more distracting than helpful, and keep us from narrowing our search for solutions before we begin.

The phrase "responsible innovation" is hardly immune from conceptual concerns. Does the adjective "responsible" imply that "innovation hitherto has been irresponsible, or at least not explicitly responsible" (Gaglio *et al.*, 2019, p. 13)? The anomaly disappears when we are reminded of RI's tendency to associate innovation with technological change,[2] which renders the term "responsible" anything but redundant given the risks posed by new technologies. But, as noted, not all innovation is technological and, as a review of the patent records will affirm, not all technological change is innovation. Thus, RI needs to be grounded on a deeper understanding of innovation – the term that "responsible" modifies. The concerns are not merely semantic: a constricted view of innovation is hardly the strongest foundation for expanding participation and exploring the full range of innovation opportunities and risks, two of RI's ambitions (Baur, 2021; Robinson, 2020).

Clearly, we need to develop a stronger account of the concept of innovation. (Blok, 2018). Although innovation could be thought of as a process (for example, von Schomberg, 2013, p. 63; Tidd & Bessant, p. 19), questions of process presuppose

[2] Von Schomberg (2013, p. 63) provided this initial definition: "Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)."

an understanding of the *goal*, which is our focus here. This article asks what criteria must be satisfied for a contribution to the world to qualify as an innovation or, more simply, what is an innovation? The term "innovation" has been appropriated for a variety of purposes through the decades (Gaglio *et al.*, 2019), and my objective is not to catalogue or reconcile the term's varied use. Instead, I will propose that we understand the term "innovation" in a particular way, roughly akin to what Stevenson called a "persuasive" definition (Boisvert, 2021). I am hopeful that the understanding I propose here will clarify and help to unify our understanding of the term and have important implications for innovation practice and for RI.

Innovation, I will contend, is *ethical change that delivers substantial applied value to beneficiaries of a domain* (hereafter, the "proposed understanding," "account" or "definition"). In the sections that follow, I will develop this account, elaborate on its elements, and explore its implications. The article will begin with a discussion of why innovation is best understood as change that makes the world *better*. Because "better" is a normative concept, I will next turn to the nature of normativity and its various dimensions. I will draw on "fitting accounts" of value to explain why value is not only a referendum on what persons desire, but also on what they are warranted in desiring. I will explain why innovation delivers *applied* values and reflect on the interpretive range of that term. The article will next consider ethics, distinguishing it from "value" and explaining why delivery of value is a necessary but not sufficient condition of innovation. Ethics, I will argue, provides more than a post-delivery critique of our innovation efforts; it is a condition of innovating in the first place.

The article will next examine the nature of innovation change, exploring the concepts of domains and beneficiaries, introducing "suboptimal states" to consider the types of value that innovators can deliver, and explaining why change that qualifies as innovation must also be substantial. Next, I will consider the challenge of "delivering" value, offering a conception of technology well beyond the highly engineered artifacts that are widely associated with the term. After developing the proposed account of innovation, I will explore its relation to the various categories of X-innovation – suggesting that the proposed account can help us unify our understanding of innovation. The article will then discuss how the proposed account can deepen and refine our understanding of RI.

A BETTER WORLD

Innovation has been a buzzword for nearly half a century (Gaglio *et al.*, 2019). One explanation for the widespread use is its association with the idea of novelty: the term innovation has been defined as a "new idea, method or device: novelty," or "the introduction of something new" (Merriam-Webster). Novelty is far from the whole

story, however. We would not characterize a change as an "innovation" if it delivered nothing of value or if it made things worse. Innovation holds a special, almost revered place in our thinking because it ranks among the most important of our human capacities. We have the power to transform our own world, and innovation is our expression of that capacity. We prize the new not as an end in itself, and not for amusement, novelty or to display our ingenuity, but for its capacity to improve our world.

In the "social" innovation setting, the notion that change should advance a social value is generally a part of the definition (e.g, Tidd & Bessant, 2018, p. 536). But in the business or industrial setting, some may view this notion as idealistic or beside the point. Businesses innovate for strategic advantage – increased market share, to establish a niche, to gain market leadership, to stave off competition – and ultimately for profit (Tidd & Bessant, 2018, pp. 9-10). Unsurprisingly, some have defined the term innovation around these or related goals.[3] Thus, some might argue that business innovation and (what is often taken as its close kin) technological innovation[4] are not motivated by a desire to improve the world but are ultimately about profit. Although we may choose to deliver change into the world for a variety of reasons, whether we have *innovated* turns not on our motive but on whether what we have delivered has changed the world for the better. Moreover, the Environmental, Social, and Corporate Governance (ESG) movement increasingly aligns shareholder investment with genuine value (Goedhart & Koller, 2020; Henderson, 2020, p. 132). Of course, some companies may nonetheless deliver change that increases competitive advantage or profit regardless of whether it improves the world.[5] That, however, is why we should reserve the term innovation for those changes – commercial or otherwise – that create value in the deepest sense.

THE NORMATIVE (PART I): VALUE

As the foregoing suggests, innovation is not simply change or novelty: it is change that improves the world. The innovator moves the world from its current state to a better state, from *is* to *ought*. As a result, innovation is inescapably *normative*, a term

---

[3] For example, "Innovation is the multi-stage process whereby organizations transform ideas into new/improved products, service or processes, in order to advance, compete and differentiate themselves successfully in their marketplace" (Baregheh 2009, p. 1334).

[4] Again, to the extent that innovation is associated with the commercialization of technological advances, the objections raised here and the responses that follow apply with equal force to "technological innovation". One understanding of the latter term emphasizes scientific knowledge that is translated into useful outcomes or products "through the marketplace" (Gaglio, Godin & Pfotenhauer 2019, p. 7).

[5] As von Schomberg and Hankins (2019) note, "RI reflects an economic paradigm that acknowledges that that market innovations do not automatically deliver on societally desirable objectives."

that "is concerned with what ought to be the case" (Wedgewood, 2010, p. 445). The normative can be parsed into four elements: the evaluative, the "reason giving," the deontic and the fitting (Cuneo, 2020). By exploring these concepts and their interrelation, we can deepen and delimit our understanding of what it means to innovate.

The evaluative is concerned with what we value or favor. Words like good, better, best are a part of our evaluative arsenal, and signal our approval. Our approval can take varied forms, however, as a taxonomy of the ways we use the term "good" reveals. *Goodness of a kind* (or attributive good) claims that something is good at what it purports to do – for example, a good chess player or a good chess clock. Alternatively, something might be *good for* another; for example, medicine is "good for" persons. *Good simpliciter*, a third variety, claims that something is good in a way that transcends their goodness "for" or "of a kind" – for example, friendship, beauty or knowledge (Schroeder, 2016, 1.1). We can also distinguish instrumental from intrinsic goods, the former valuable for what they bring about, and the latter for their "own sake" (Rowland, 2015, p. 203).

Although the terms "value" and "good" are often used interchangeably, here I will generally use the term "value" or "values". For example, when we say that medicine is "good for" a person, we are implicitly claiming that it provides something of value for that person – in this case, health. Likewise, with goodness of a kind: when we say that Louise is a good artist, we are makings claims about how her work aligns with what we value in paintings – a certain handling of light and perspective, for example. I will refer to the variety of things that we favor in any setting as the "values" in that setting.[6] This will enable us to itemize and compare what is worthy of approval without referring in each instance to "things that are of value" or a similar formulation. The use of the term "values" rather than "goods" will also help avoid confusion with the everyday use of the term "goods" as "items for sale" (Cambridge Dictionary). In using the term "value", however, I am not referring to the *degree* of worth we ascribe or to market or exchange value.

Our next question is whether the good or values are anything more than subjective? A "dispositional" view holds that the good is whatever normal persons in normal circumstances deem it. When we characterize something as good, however, there is a sense in which we are offering more than a report on the reaction of others. The strongest account of this "more" is "robust realism," which holds that "values exist independently of human responses to them" (Jacobson, 2011, Introduction). This view

[6] A similar use of the term "values" can be found in value-sensitive design, where it is understood that designs can "produce or reproduce" a variety of "values" (Dignum *et al.*, 2016).

strays too far from the practical goals of innovation, however, by rendering irrelevant the response and sensibility of the innovation's intended beneficiaries.

The "fitting attitude" approach fuses both standards and provides the best account of the value that innovators must deliver. This approach insists on both beneficiary approval and that their approval be fitting – that, is warranted, correct, appropriate, or the like. Critically, however, a fit attitude is not a *moral* claim; it is a claim about whether our approval is appropriate. Fitness is what we "ought" to value (Jacobson, 2011, Section 1). Suppose, for example, that Mary's fans deem her a good singer even though she is invariably off key. In suggesting otherwise, we are not claiming that her fans' approval is immoral, but that it is unwarranted. Thus, on the account proposed here, change that meets with the approval of those to whom it is delivered would qualify as valuable only if that approval is also fitting.

Philosophers have offered rich alternative conceptions of what makes an attitude fitting. One widely discussed account of the good has been coined the "buck passing" account by T.M. Scanlon. It holds that for something to be good (or bad) is for it to have properties that constitute reasons to have pro (or anti) attitudes towards it (Scanlon, 1998, Chapter 3). The buck-passing account has arguably emerged as the majority view that "varieties of value, and in fact other properties such as wrongness, oughts, and fittingness, should all be understood in terms of normative reasons for pro-attitudes or actions" (Orsi, 2020, p. 653). Although the "reasons" that are cited are natural or psychological facts about the world, whether they warrant the conclusion of good or valuable has an "open feel" (Scanlon, 1998, p. 96).

That open feel calls on our practical judgment about the reasons on offer – even for goodness of a kind, where the kind seemingly supplies its own standards. Scanlon writes:

> [F]or example, a good thermometer might be thought to be one that controls a furnace in such a way as to maintain a set temperature. This would simply be a claim of physical fact. But something would not be a good thermostat if it were the size of the Empire State Building, or took as much energy to operate as the furnace itself. The ''purposes'' or ''interests'' relative to which we judge something a good thermostat include a variety of more specific considerations, and such a conclusion about goodness requires a judgment about the proper balance between these considerations. (Scanlon, 2011, p. 446)

Innovation often dwells in these "more specific considerations" and the practical judgment about their value. What Scanlon calls "purposes or interests" would be "values" on our schema. As an agent of change, the innovator invites us to reconsider how well these values are delivered, and perhaps to reset the balance between them, or even expand the variety of values associated with the "kind" in question.

What makes a thermostat good? In the decade since Scanlon penned those words, innovators have delivered thermostats that: determine whether we have left

the home, adjusting the temperature accordingly; use machine learning to make inferences about our temperature preferences; can be set remotely; and rethink the user interface and physical appearance of the device. These various technologies (a term we will explore below) each purport to deliver some value: remote sensors gather information that will help deliver *location-specific temperatures*; machine learning delivers *time saving, convenience, reduced energy consumption, cost savings*; and design delivers *elegance and ease of use*. Perhaps all of these values were a part of our earlier understanding of the thermostat, and the innovator has simply found ways to deliver them substantially better. On the other hand, the aesthetic appeal of the device could be understood as a *new* value for those who have never reflected on appearance in a thermostat.

As our thermostat example suggests, a change can improve on the delivery of multiple values. It is also possible that a change could implicate a choice *between* values – for example, a device that saves time by translating voice into commands but collects user communications to enhance performance. Assuming that the privacy concern is better understood as a disvalue than as an ethical breach, the question is by what criterion we can evaluate the proposed device given that convenience and privacy are arguably incommensurable – that is, they share no common scale for assessment. On some accounts, however, we can and do make choices between incommensurable values, and do so for *reasons* (Sunstein, 1994, p. 809-810). While the particulars of these accounts are outside our scope, I will assume here that the complex and potentially incommensurable range of values implicated by a change does not foreclose an assessment of whether they deliver substantial improvement.

"Value" on the proposed account includes whatever we fittingly favor. Innovation is change that improves the world, and our understanding of value is as expansive as that notion. That, in turn, suggests that the commonplace association of innovation with commercial or marketplace efforts noted by von Schomberg and Blok (2018, p. 6) vastly underestimates the varieties of value that innovators can deliver. Moreover, while market acceptance certainly amounts to approval, we must also ask whether that approval is fitting.

Innovation differs from a great new idea, theory, explanation, symphony or work of art. Scientific breakthroughs and artistic masterworks are profoundly valuable in their own right; in fact, knowledge and beauty are widely understood as basic goods (for example, Finnis, 1980). Despite their profound importance, however, we have more apt terms for those basic contributions. We consider Einstein's relativity theory a seminal advance in our knowledge, not an "innovation." By contrast, we consider GPS navigation devices an "innovation," even though the understanding provided by Einstein's relativity theory was essential in making the devices accurate (Dijkgraaf,

2017). Innovation takes place at the applied level, where we hope to effect practical, palpable improvements in the lived world – and thus the proposed definition insists on *applied* value.[7]

Of course, there is ample room for disagreement about whether a value is "applied." For example, we might consider a vast improvement of the beauty and elegance of a device an "innovation," while the term seems inapt (and insufficient) to describe an artistic masterpiece. To complicate matters further, technological innovations often power scientific advances (in addition to the reverse path described above). And, of course, knowledge and applied advances are often inextricably linked: for example, Claude Shannon's application of Boolean algebra to computation and his theory of information advanced our understanding at a fundamental level while enabling radical advances in computation and data transmission, respectively (Soni & Goodman, 2017). Despite the healthy room for interpretation, however, an emphasis on "applied" values helps to cabin the concept of innovation by emphasizing the practical improvement it seeks to deliver.

THE NORMATIVE (PART II): THE ETHICAL

Another element of the normative domain is the moral, which addresses "ought" not in the sense of what we ought to favor, but in terms of duties owed to others. Regardless of one's position on the relationship between the good and ethical, I will contend here that innovators must not assume alignment between change that delivers value and its ethicality. Innovation, as argued earlier, requires change for the better. The implications are twofold: innovation entails the *ethical* delivery of value; and value and ethics are distinct inquiries.

Agents of change often work to deliver values that are best understood as "goodness of a kind" or "goodness for" others, categories explored earlier. It is important to distinguish this goodness from an ethical claim. For example, the "good" in goodness of a kind is not the greater good or an increase in general welfare but a standard based on the values addressed by the object in question. Imagine an encryption tool that is nearly unbreakable. Surely, the tool is good at what it purports to do and might be fittingly favored or be said to deliver value for that reason. Some have raised ethical concerns about these tools, however, because they may facilitate unlawful behavior (Bay, 2017). These ethical concerns are categorically distinct,

[7] Andrew Maynard defines innovation as "the translation of creative ideas into products and processes which provide sufficient value to others that they are willing to invest in them." He emphasizes the applied, practical nature of innovation, which is "focused, targeted, purposeful change, rather than undisciplined creativity and undirected invention" (2020, p. 118).

however, from the tool's goodness of a kind: the encryption tool poses ethical concerns, on this critique, *because* it does its job so well.

One could argue that we should redefine the purpose of the object to include ethical criteria in assessing "goodness of a kind" or "good for". But a sweater is not good because it purports to increase general welfare or satisfy deontological concerns but because it satisfies the criteria or values of its kind – say, its warmth, attractiveness, comfort, ease of cleaning, and so on. If the sweater is made under unfair labor conditions, we would not say that those workers made a "bad" sweater, but that the sweater (however good) was unethically sourced. A similar concern attends the category of "goodness for". The claim of value is directed at a beneficiary and does not purport to address the world as a whole. As a result, a claim of "goodness for" makes no claim of alignment with ethical considerations.

Agents of change might mistakenly assume that the creation of value is the sole measure of whether they have made the world a better place. A change might deliver benefits or "values" to some, however, without addressing the harms created in delivering that value. Ethical inquiry, by contrast, is not limited to those who benefit from the change; it can be understood as the duties we owe to others – *tout court*. As a result, innovation must deliver value *and* do so ethically. Simply put, change that is not ethical is not innovation. And it should not be delivered into the world.

THE VARIETIES OF ETHICAL, VALUABLE CHANGE

We can now develop a simple taxonomy what constitutes valuable, ethical change. We will begin by defining terms. Change implies an initial and new state of affairs. That, in turn, implies a *domain* in which the initial and new state of affairs occur. Domains are the artifacts, activities, institutions, groups, communities, regions, or any other setting or category that is the subject of the innovator's efforts to deliver change. Domains can range from the local to the universal, and whether one has innovated does not depend on the size or reach of the domain. The change within those domains must deliver value, however, and therefore innovation domains have *beneficiaries*, i.e., those who will benefit from the change.[8] We will refer to those who are not the intended beneficiaries of the change but are nonetheless affected by it as "third parties."

[8] The would-be innovator's proposal may or may not be accepted by the beneficiaries. In order to constitute an innovation, however, the change must, inter alia, be embraced by the beneficiaries and their approval must be fitting or warranted.

Now that we have introduced the notion of domains and beneficiaries, we can examine the idea of *valuable change* more closely. A doctor who routinely performs a lifesaving surgery is delivering profoundly valuable outcomes, but we would not consider her surgical efforts an innovation. By contrast, we would consider the lifesaving surgical *method* she deployed an innovation when it was introduced. As von Schomberg and Blok (2018, p. 9) note, "innovation does not refer to the simple introduction of new music but to the introduction of a new way of making music." When we use the domain's pre-existing knowledge and methods to deliver valuable outcomes for others, we are not introducing change into the domain, we are practicing its methods. The surgical method, by contrast, delivers a change into the domain, and change of this kind would constitute innovation if the other elements of the definition are satisfied.

Change can deliver degrees of improvement, which raises the question of whether all improvements, no matter how modest, qualify as innovation. An expansive understanding in part explains the use by some of qualifiers such as "routine" and "radical" to differentiate innovations (Pisano, 2019, p. 31). But if every improvement, no matter how slight or incremental, is considered innovation, we lose sight of why innovation holds a revered place in our thinking. Innovation promises to improve our lives substantially, and this is why we are committed to learning its methods and practicing its art. Admittedly, by reserving the term innovation for change that delivers "substantial" value, the normative challenge is redoubled – we must now reflect on questions of value *and* degree. But innovation is ultimately a normative enterprise, and these questions are best embraced rather than defined away for want of simplicity or algorithmic answers.

The innovator therefore moves us from an initial state to a substantially improved state within a domain – from "is" to "ought". The would-be innovator sees the initial state as suboptimal (which I will term a "suboptimal state" or SOS), where "optimal" signifies what might within practical reason be addressed rather than the ideal. Of course, we typically do not know in advance whether the gap can be closed; as a result, the "ought" in the suboptimal state frames the target for the innovation journey. Innovation on this account is thus the curing of a suboptimal state.

Suboptimal states can take two forms – "functional" and "value". A functional SOS is a substantial gap between a *current* value in the domain and its optimal state. It is a claim, in other words, that one or more of its current values can be delivered substantially better; the gap is in the *functionality* not the value. This is how we might understand Atul Gawande's introduction of the checklist into the surgical rooms of developing countries, radically improving medical outcomes (Gawande, 2010). The value – healthy surgical outcomes – was already a goal of this domain. The critical

change was in introducing a method into the domain to substantially improve *delivery* of that value.

Functional gaps are often the easiest to recognize because we are quickly frustrated when our expectations are not met. Suboptimal states are not modest deficiencies, however, or their resolution would lead only to incremental change. For example, software users are quick to identify ways in which a program might operate a bit more effectively. But solving programming bugs or adding minor program features is not innovation. By contrast, while cars have long sought to keep their occupants safe, the annual death and injury statistics suggest a substantial gap in delivering that basic value. If self-driving cars substantially reduce auto deaths and injuries, they will resolve a functional SOS, and qualify as an innovation.

A *value* suboptimal state, by contrast, is a gap between the current values of a domain and those important values one suspects it *could* deliver. The value SOS is not resolved by the improved delivery of a current value; it is the delivery of an important new value or value emphasis within the domain. For example, Tesla made the environmental benefits of its electric automobile a central part of its story in addition to the car's performance on more traditional automobile metrics (Tesla, 2019). When restaurants began emphasizing locally grown food on their menus, they presented a distinctive and new value for customers. Sometimes, new values will attract new beneficiaries to a domain: for example, some libraries have re-envisioned their role as helping the members of the community to build skills. Once again, however, we are remitted to judgment in determining whether the change qualifies as an innovation: the new value or value emphasis must be important if we are to distinguish incremental improvements from innovation.

A value suboptimal state can also be understood by what we hope to *eliminate* from a domain, i.e., a "disvalue." Here, we are not speaking of new values that we would like to introduce into the domain, but the presence of something we have reason to disfavor and therefore reduce or eliminate. For example, if machine learning delivers biased decision-making, that bias is a disvalue that an innovator might seek to eliminate. For simplicity, we will treat the addition of value and the elimination of disvalue as the delivery of value to a beneficiary in our definition.

For all its importance, the delivery of substantial, fitting value is a necessary but not a sufficient condition of innovation. Change that delivers value, no matter how substantial, is not innovation if it treats the beneficiaries or third parties unethically. As a result, ethics is not solely an after-the-fact constraint on innovation efforts; it is a critical element of the innovation process itself. By addressing both the value and ethical dimensions in the innovation process, the innovator reduces harm and the

need for subsequent interventions that can prove more costly and challenging after the fact.

The proposed understanding brings into bold relief the potentially fraught relation between values and ethics for the innovator. Changes often deliver substantial value to beneficiaries while posing risks to beneficiaries who *may* have consented and third parties who likely have not. The proposed understanding encourages innovators and the public to see these not simply as "benefits and costs" questions but as "values" and "ethics" questions, inviting a richer discussion of whether the values in question are substantial and fitting while accommodating utilitarian, deontological (and other) ethical assessments of whether and how they should be delivered.

THE DELIVERY OF VALUE – TECHNOLOGY RECONCEIVED

As we have seen, valuable change can take the form of substantially improved delivery of a current value within a domain or the delivery of an important new or revised value into the domain.[9] In either instance, the innovator must find a way to *deliver* something of value into the domain. The term "deliver" is intended to evoke the practical, real world outcomes expected of the innovator: he/she must bring about or effect a palpable change in the domain. The proposed understanding does not identify or prioritize any particular *type* of delivery, however; instead, it seeks to accommodate the vast and varied means of delivering value. This understanding, in turn, has important implications for our understanding of the term "technology" – a term so widely associated with innovation (Gaglio *et al.*, 2019) that it warrants closer inspection here. How *does* technology relate to innovation, and equally important, how should an innovator understand that term?[10]

We can begin with Brian Arthur's definition of technology as, "A means to fulfill a purpose: a device, or method, or process" (2009, p. 29).[11] To refine that formulation, we might add the term "tools" which we tend to see as distinct from devices. Because "process" and "method" by and large capture the same notion, we will use only the latter term. We might add "materials," since they too deliver functionality and are commonly distinguished from means and tools in everyday parlance (for example,

[9] This discussion also applies to the substantial reduction of disvalue in a domain, since it too requires some means of effecting the chosen outcome.

[10] The nature of technology has been the subject of inquiry from a variety of disciplinary perspectives, such as philosophy, sociology, and engineering. Here, the question is directed only at how we might understand technology in light of the proposed understanding of innovation.

[11] Pitt (2000) also offers an instrumental account of technology, characterizing it in a shorthand account as "humanity at work" or more formally as "the deliberate design and manufacture of the means to manipulate the environment to meet humanity's changing needs and goals" (p. 30-31).

Brownell, 2017; Tibbits, 2021). Of course, these terms overlap: tools and materials instantiate methods; and devices, methods and materials could be understood as tools, since they seek to effect some purpose or goal. Nonetheless, the categories provide a taxonomy of the ways that technology enables us to innovate within a domain. Thus, from an innovator's perspective, technology might be understood as "tools, devices, methods and materials" that deliver something of value into the world.[12]

This understanding of technology suggests the vast range of sources that innovators can draw on in delivering change into a domain. Many associate innovation with highly complex technologies such as machine learning, blockchain or the gene-editing tool, CRISPR. These are immensely powerful tools that auger vast and valuable change. But simpler methods also have stunning change power when applied in the right setting. Consider the simple checklist mentioned earlier that Atul Gawande introduced into the surgical rooms of the developing world, substantially improving outcomes; or reforestation to reduce global warming (de Groot, 2019).

Von Schomberg and Blok (2018) have documented the widespread association of innovation with emerging technologies. That belief has been reinforced by the extension of Stuart Kaufman's concept of the "adjacent possible" to the innovation setting, whereby "[e]ach new combination opens up the possibility of other new combinations" (Planing, 2017). While emerging technologies are profoundly important,[13] they hardly exhaust the means of delivering values – even for technology as traditionally understood. Consider "exaptation" or the repurposing of technologies to create valuable change. Gutenberg's repurposing of the wine press into a printing press offers a seminal example. As Johnson (2010, p. 153) notes, "An important part of Gutenberg's genius… lay not in conceiving an entirely new technology from scratch, but instead from borrowing a mature technology from an entirely different field, and putting it to work to solve an unrelated problem".

---

[12] One might question this instrumentalist understanding of technology given the various accounts of technology that emphasize its human experience, ethical or value implications (for example, Verbeek, 2006; Winner, 1978). The proposed understanding does not deny the claim that technology once deployed is fraught with these implications; it is premised on that claim. Technology, as understood here, is the means to deliver normative outcomes. In seeking to cure suboptimal states, however, the innovator must attempt to decouple the instrumental elements of a technology from its currently deployed normative elements if only to ask whether it can ethically deliver the values that she hopes to deliver. Once the innovator has forged a connection between means (technology) and ends (values), she must engage in the analysis of its value and ethical implications before delivery into a domain, an inquiry that can be deepened by RI's commitment to anticipatory and reflexive innovation of RI (for example, Stilgoe *et al.*, 2013) and value-sensitive design's efforts to "intentionally embed desired values into technologies" (Simon, 2017).

[13] Some new technology may advance the state of art and qualify as an invention, but inventions may or may not deliver the substantial value required of innovation, as the patent rolls will affirm.

Repurposing, in turn, suggests a distinction between *technology advances* and *advanced technology*. Expertise is typically necessary to advance the state of art by developing new and complex technology. Conversely, the capacity to reflect on domains and values may prove more important than expertise in seeking ways to repurpose technology that is already developed – even when it is advanced. As discussed earlier, whether an innovator is addressing a functional or value gap, the degree of change that ultimately matters is in the *value state* of the domain, a change that may or may not require substantial technological advances.

Technology for the innovator also includes methods that are not traditionally associated with the term, such as ways to structure social settings to deliver values, roughly akin to the expansive understanding proposed by Pitt (2000). For example, Jane Jacob's proposal to make urban neighborhoods safer – such as mixed uses, and people and eyes on the street (Jacobs, 1961) – could be understood as a technology for an innovator, like any other method that delivers values into world. So too could the subscription business model that enables some local, organic farms to survive (Neumark, 2017); the "nudge" of changing the default from opt-in to opt-out to, among other things, increase retirement investment (Thaler & Sunstein, 2008); or a ranked-choice voting method that reduces polarity (Kambhampaty, 2019). While this more expansive understanding of technology does not accord with common usage, it invites would-be innovators to draw on the widest array of means to deliver values into a domain.

A broad understanding of technology may help to correct the tendency of some to favor highly engineered solutions either without warrant or without addressing the cultural and social barriers to their adoption (Toyama, 2015). When we understand that behavioral, sociological, and cultural methods (to name only a few) are also "technologies" in the sense that they deliver value into domains, we are more likely to draw on the right means or combinations of means in crafting solutions. New, complex, and highly engineered artifacts are profoundly important, of course, but they are one of the many ways to deliver change. Innovators should draw on a vast palette of tools, methods, devices and materials – from the simple to the complex, to the old and new – to deliver valuable change.

RELATION TO X-INNOVATION AND RI

A profusion of adjectives attempts to categorize innovation paths. Terms like "technological," "industrial," "social," "open" or "sustainable" innovation have prompted Gaglio *et al.* (2019) to describe the phenomenon as "X- innovation." Although these categories can help organize efforts around shared methods and goals, they also pose concerns. If, as argued here, improving the world is the *raison detre* of innovation, then

our efforts may be best oriented around questions of *value*, which may lead us across business, social, technology, and other boundaries. Likewise, while funding to develop and sustain any innovation is essential, what funding mechanism is best – commercial, nonprofit, public support, donation, or other – may *depend* on the innovation breakthrough rather than *inspire* it. Likewise, the emphasis on technology innovation, while often helpful as an organizing principle, risks prioritizing the selection of means over goals, which can narrow the range of value delivered. The proposed understanding invites us instead to see innovation not as a variety of predetermined paths, but as the movement toward substantial value – with funding and means as elements that emerge in time.

Responsible Innovation differs from other forms of X-innovation. It does not prescribe a singular path, but instead can be understood as a framework to promote "collective stewardship" of innovation (Stilgoe *et al.*, 2013, p. 1570). The term "responsible" has varied meanings, but the dictionary definitions "trustworthy," "sensible," "morally principled," and "ethical" (Oxford English Dictionary) suggest something of its intended role in RI. A critical question, therefore, is why qualify the term "innovation" with "responsible" if "innovation" (as understood here) already incorporates these normative notions? One explanation, of course, is that we have long struggled to agree on a definition of innovation and, as discussed earlier, many associate innovation with change – often technological change – decoupled from normative elements. The "responsibility" qualifier is warranted if we see innovation through this narrow gaze.

But how are we to understand RI if we accept the understanding of innovation proposed here? The proposed understanding orients innovation around the ethical delivery of substantial (and fitting) value but it does not tell us *how* to accomplish this. RI's vast collection of practices – including, for example, anticipation, reflection, participation, and responsiveness (Stilgoe, Owen & Macnaghten 2013) – can be understood as the means to satisfy the normative elements of innovation. Understood thus, RI's importance stems not from the misconception that innovation, *per se*, is agnostic on questions of value and ethics, but *because* innovation, properly understood, entails these commitments. RI offers a framework for implementing this understanding.

Moreover, the proposed understanding can provide theoretical grounding for RI's varied practices and continued development. For example, RI "places a premium on inclusive participation that allows the setting of research and innovation goals" (Owen & Stigloe 2012, p. 754). The proposed understanding helps explain why the discovery of "suboptimal states" (and therefore innovation goals) is the province of everyone who is attuned to the workings of a domain and why those goals need not

be solely commercial. In addition, the proposed understanding holds that change must ethically deliver fitting value in order to *qualify* as innovation, which helps explain RI's emphasis on incorporating its responsible practices *into* the innovation process rather than relegating them to post-delivery critique (Bauer *et al.*, 2021).

Third, RI envisions that the public serve as an "active player that can contribute with innovative ideas" (Robinson *et al.*, 2020, p. 3). Under the proposed understanding, the functionality that delivers value into a domain need not be new or complex, and the repurposing of even complex technologies requires far less expertise than the underlying advances that made them possible. In addition, innovation may result from the delivery of new or remixed values into a domain rather than new or complex technology. Thus, the understanding proposed here supports RI's expansive vision who can and should *innovate*.

Owen and Pansera have observed that "most academics working in the field [see RI's initiatives as]" a set of policy agendas and action lines which structure a work program of the EC, rather than a coherent and intellectually robust discourse" (2019, p. 38). It is, of course, difficult to build a conceptual understanding of RI without an understanding of the concept of innovation. Once we embrace the normative reach of innovation, we can begin to build conceptual bridges to RI's varied commitments. This may prove an important line for future scholarly work.

CONCLUSION

Innovation is inescapably a normative enterprise, and we can construct an understanding of the term by exploring its normative elements. Innovation is best understood as ethical change that delivers substantial applied value to beneficiaries of a domain. Value on this understanding is not only what we favor but what we have *reason* to favor. Innovators change the value state of a domain, moving it from is to ought. They cure "suboptimal states" by substantially improving the delivery of current values in a domain or by delivering important new or remixed values into a domain. But the delivery of substantial applied value is not enough. *Ethical* delivery of that value is an additional condition of innovation, requiring the innovator to address the risks posed to beneficiaries and third parties by the proposed change.

This value-centered understanding of innovation suggests much about the nature and role of technology. For the innovator, technology is best understood as *any* tool, device, methodor material that delivers value. The functionality that delivers value can range from the simple to the complex, the old to the new.

What's in a definition? A great deal it seems. If we see innovation as simply novelty or change, we miss the principal reason for innovating – to improve the world.

When we see innovation as only what creates market value, we conflate approval with fitting approval, dismiss the vast array of value that may not be captured through exchange mechanisms, and forget that value and ethics, *together*, are the *sine qua non* of innovation. If we see innovation as principally the product of highly complex technology advances, we overlook myriad other ways to deliver valuable change and unduly limit the values we might deliver.

The understanding proposed here invites us to unify our understanding of innovation rather than characterizing it by method or domain. When instead we see innovation through social, technological, business, or similar lenses, we may find that our change efforts refuse to respect our prefigured categories. While these categories offer administrative guidance, they risk limiting our search for suboptimal states and the means to solve them before we begin.

We can also deepen and refine our understanding of Responsible Innovation. The proposed understanding grounds our understanding of innovation in its normativity. RI can be understood, in turn, as an effort to operationalize those normative elements: it can help us determine the values worth pursuing, the ethical risks to be addressed, and the means by which they are pursued. Moreover, the broad understanding of change methods described here can broaden RI's approach to innovation and who is seen as a prospective innovator. The foundation for a sound conceptual framing of Responsible Innovation lies in the normative nature of its object – innovation.

REFERENCES

Arthur, W. B. (2009). *The Nature of Technology: What It Is and How It Evolves*. Free Press.

Baregheh, A., Rowley, J., & Sambrook, S. (2009). Towards a multidisciplinary definition of innovation. *Management Decision*, 47(8), 1323-1339. https://doi.org/10.1108/00251740910984578

Bauer, A., Bogner, A., & Fuchs, D. (2021). Rethinking societal engagement under the heading of Responsible Research and Innovation: (novel) requirements and challenges. *Journal of Responsible Innovation*, 8(3), 343-363. https://doi.org/10.1080/23299460.2021.1909812

Bay, M. (2017). The ethics of unbreakable encryption: Rawlsian privacy and the San Bernardino iPhone. *First Monday,* 22(2). https://doi.org/10.5210/fm.v22i2.7006

Blok, V. (2018). Philosophy of Innovation: A Research Agenda: Guest Editorial. *Philosophy of Management*, 17(1), 1-5. https://doi.org/10.1007/s40926-017-0080-z

Blok, V. (2021). What Is Innovation? Laying the Ground for a Philosophy of Innovation. *Techné: Research in Philosophy and Technology*, 25(1), 72-96. https://doi.org/10.5840/techne2020109129

Boisvert, D. R. (2021). Charles Leslie Stevenson. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Fall 2021 Edition). Stanford University. https://plato.stanford.edu/entries/stevenson/

Brownell, B. (2017). *Transmaterial Next: A Catalog of Materials that Redefine Our Future*. Princeton Architectural Press.

Cambridge Dictionary (n.d.). Goods. In *Cambridge Academic Content Dictionary*. Cambridge University Press. https://dictionary.cambridge.org/us/dictionary/english/goods

Cuneo, T. (2020). Normative. In *A Dictionary of Ethics*. Oxford University Press. https://doi.org/10.1093/acref/9780191835759.001.0001

de Groot, H. (2019). A Low-Tech Climate Fix. *Scientific American*, 320(4), p. 7.

Dignum, M., Correljé, A., Cuppen, E., Pesch, U., & Taebi, B. (2016). Contested Technologies and Design for Values: The Case of Shale Gas. *Science and Engineering Ethics*, 22(4), 1171-1191. https://doi.org/10.1007/s11948-015-9685-6

Dijkgraaf, R. (2017). The World of Tomorrow. In A. Flexner, *The Usefulness of Useless Knowledge* (pp. 3-47). Princeton University Press.

Finnis, J. (1980). *Natural Law and Natural Rights*. Oxford University Press.

Gaglio, G., Godin, B., & Pfotenhauer, S. (2019). X - Innovation: Re-Inventing Innovation Again and Again. *NOvation-Critical Studies of Innovation*, 1(June), 1-17.

Gawande, A. (2010). *The Checklist Manifesto: How to Get Things Right.* Picador.

Goedhart, M., & Koller, T. (2020). The value of value creation. *McKinsey Quarterly,* June 16. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-value-of-value-creation.

Henderson, R. (2020). *Reimagining Capitalism in a World on Fire*. Public Affairs.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. Random House.

Jacobson, D. (2011). Fitting Attitude Theories of Value. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 2011 ed.). Stanford University. https://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/.

Johnson, M. (2018). *Reinvent Your Business Model: How to Seize the White Space for Transformative Growth*. Harvard Business Review Press.

Johnson, S. (2010). *Where Good Ideas Come From: The Natural History of Innovation*. Riverhead Books.

Kambhampaty, A. P. (2019). New York City Voters Just Adopted Ranked-Choice Voting in Elections. Here's How It Works. *Time*, November 6. https://time.com/5718941/ranked-choice-voting/

Lubberink, R., Blok V., Van Ophem, J., & Omta, O. (2017). Lessons for Responsible Innovation in the Business Context: A Systematic Literature Review of Responsible, Social and Sustainable Innovation Practices. *Sustainability*, 9(5), 721-756. https://doi.org/10.3390/su9050721

Maynard, A. (2020). *Future Rising: A Journey from the Past to the Edge of Tomorrow*. Mango Publishing.

Merck (2021). Over 30 Years: The Mectizan® Donation Program. *Responsibility*, January 6 https://www.merck.com/stories/mectizan/

Merriam-Webster (n.d.). Innovation. In *Merriam-Webster.com dictionary*. Merriam-Webster. https://www.merriam-webster.com/dictionary/innovation

Neumark, L. (2017). How the CSA Model Supports a Farm. *HuffPost,* December 6. https://www.huffpost.com/entry/how-the-csa-model-support_b_6697734

Orsi, F. (2020). The Normative and the Evaluative. The Buck-Passing Account of Value. *The Philosophical Quarterly*, 70(280), 652-655. https://doi.org/10.1093/pq/pqz075

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: from Science in Society to Science for Society, with Society. *Science and Public Policy*, 39(6), 751-760. https://doi.org/10.1093/scipol/scs093

Owen, R., & Pansera, M. (2019). Responsible Innovation: process and politics. In R. von Schomberg & J. Hankins (Eds.) *International Handbook on Responsible Innovation: A Global Resource* (pp. 35-48). Edward Elgar.

OED (n.d.). Responsible. Oxford English Dictionary.

Parkes, C. (2008). Introduction. *OMEGA – Journal of Death and Dying*, 56(1), 1-5.

Pisano, G. P. (2019). *Creative Construction: The DNA of Sustained Innovation*. Public Affairs.

Pitt, J. (2000). *Thinking About Technology*. Seven Bridges Press.

Planing, P. (2017). On the origin of innovations – the opportunity vacuum as a conceptual model for the explanation of innovation. *Journal of Innovation and Entrepreneurship*, 6, 5. https://doi.org/10.1186/s13731-017-0063-2

Robinson, D., Simone A., & Mazzonetto, M. (2020). RRI legacies: co-creation for responsible, equitable and fair innovation in Horizon Europe. *Journal of Responsible Innovation*, 8(2), 209-216. https://doi.org/10.1080/23299460.2020.1842633

Rowland, R. (2015). Reasons As The Unity Among The Varieties Of Goodness. *Pacific Philosophical Quarterly*, 97(2), 200-227. https://doi.org/10.1111/papq.12057

Scanlon, T. M. (1998). *What We Owe To Each Other*. Cambridge: Belknap Press.

Scanlon, T. M. (2011). The Unity of the Normative. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 154(3), 443-450. https://doi.org/10.1007/s11098-011-9736-z

Schroeder, M. (2016). Value Theory. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Fall 2016 ed.). Stanford University. https://plato.stanford.edu/archives/fall2016/entries/value-theory/

Simon, J. (2017). Value-Sensitive Design and Responsible Research and Innovation. In S. Hansson (Ed.), *The Ethics of Technology: Methods and Approaches* (pp. 219-235). Rowman & Littlefield.

Soni, J., & Goodman, R. (2017). *A Mind at Play: How Claude Shannon Invented the Information Age*. Simon and Shuster.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a Framework for Responsible Innovation. *Research Policy*, 42(9), 1568-1580. https://doi.org/10.1016/j.respol.2013.05.008

Sunstein, C. R. (1994). Incommensurability and Valuation in Law. *Michigan Law Review*, 92(4), 779-861. https://doi.org/10.2307/1289693

Telsa (2019). *Impact Report 2019*. https://www.tesla.com/ns_videos/2019-tesla-impact-report.pdf

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Tibbits, S. (2021). *Things Fall Together: A Guide to the New Materials Revolution*. Princeton University Press.

Tidd, J., & Bessant, J. (2018) *Managing Innovation: Integrating Technological, Market and Organizational Change* (6th ed.). John Wiley.

Toyama, K. (2015). *Geek Heresy: Rescuing Social Change from the Cult of Technology*. Public Affairs.

Verbeek, P. (2006). Materialising Morality: Designing Ethics and Technological Mediation. *Science, Technology and Human Values*, 31(3), 361-380. https://doi.org/10.1177/0162243905285847

Von Schomberg, L., & Blok, V. (2018). The turbulent age of innovation. *Synthese*, 198, 4667-4683. https://doi.org/10.1007/s11229-018-01950-8

Von Schomberg, R. (2013). A vision of responsible innovation. In R. Owen, M. Heintz, M. & J. Bessant (Eds.), *Responsible Innovation: Managing the responsible innovation of science and innovation in society* (pp. 51-74). John Wiley.

Von Schomberg, R., & Hankins, J. (2019). Introduction to the International Handbook on Responsible Innovation. In R. von Schomberg & J. Hankins (Eds.) *International Handbook on Responsible Innovation: A Global Resource* (pp. 1-11). Edward Elgar.

Wedgwood, R. (2010). The Nature of Normativity: Précis. *Philosophical studies*, 151(3), 445-448. https://doi.org/10.1007/s11098-010-9553-9

Winner, L. (1978). *Autonomous Technology. Technics Out of Control as a Theme in Political Thought*. MIT Press.

Is Responsible Innovation Possible? The Problem of Depoliticization for a Normative Framework of RI

Lisann Penttilä*

* KU Leuven

ABSTRACT

Global frameworks of RI face several challenges. These include (1) the extent to which economic interests can be reconciled with the concept of responsibility in innovation processes and (2) the lack of a strong political dimension in these frameworks' conception of responsibility. Responding to these challenges is crucial if RI is to ensure that societal and natural needs are sufficiently considered in the innovation process. The influence of economic incentives on innovation processes coupled with the widespread depoliticization of persons makes it more urgent that RI adopt a *political* conception of responsibility in order to safeguard the legitimacy of the values and outcomes it deems societally desirable. This paper argues that the operative conception of responsibility in RI should be broadened to include a stronger political dimension. To this end, I turn to the work of Hannah Arendt to continue rethinking the concept of "the political" and, by extension, how responsibility and politics can be understood as two sides of the same coin.[1]

Keywords: Responsibility; Politics; Depoliticization; (R)RI; Hannah Arendt.

Proposal Submitted 4 October 2020, Article Received 6 May 2021, Reviews Delivered 5 June 2021, Revised 31 October 2021, Accepted 8 November 2021, Available online 28. February 2022.

[1] Acknowledgement: I would like to express my gratitude to the reviewers for their insightful and constructive feedback, as well as Blake Scott for assisting me with the editing process.

INTRODUCTION

Innovation, by virtue of its character of novelty, confronts us with the unfamiliar and the unpredictable. While the concept of innovation has a long history, it is not until the 20th century, that it begins to develop its specifically technological and commercial connotations (Godin, 2015; von Schomberg & Blok, 2019). Alongside this conceptual development, a new generation of technology began to emerge – consider for instance the developments of AI, nanotechnology, and digital technology. Technological innovations have had both positive and negative consequences. The desire to steer innovation processes in the "right" direction and deal with the unpredictability inherent in innovation, has prompted the now vast literature of Responsible Innovation (RI). This has become important for the bodies of scientific governance that try to respond to the negative impacts understood to be the consequence of past scientific and technological innovation – an example of this is the normative "Responsible Research and Innovation" (RRI)[2] framework project of the European Commission (European Commission, 2014). However, there are conceptual shortcomings with the concept of Responsible Innovation (RI). For example, innovation in the context of RI is unreflectively understood as technological (von Schomberg & Blok, 2019), inherently good, and viewed from an economic vantagepoint (Blok & Lemmens, 2015). At the same time, the concept of responsibility in RI finds itself in the midst of interests – e.g., moral, political, and environmental – that are difficult to reconcile with technological and market-interests (Grunwald, 2018). In turn, this problematizes the legitimacy of the RI framework.

In this paper I aim to contribute to the literature that seeks to politicize responsible innovation (van Oudheusden, 2014; Owen & Pansera, 2019). To do so I first outline leading RI proponent Renè von Schomberg's critical evaluation of the hinderances facing RI today and what he takes to be central departure points for a vision of responsible innovation. Second, I argue that due to RI's insufficiently political conception of responsibility, it struggles to address the depoliticization of persons and societies which problematizes its own responsibility agenda.[3] To address this, in a third step I engage with the work of philosopher and political theorist Hannah Arendt to shed light on how responsibility and politics can be understood as two sides of the same coin. In turn, this can allow us to see in what sense depoliticization problematizes the "responsible" and normative agenda of RI. In the fourth and final

[2] The abbreviations RI and RRI are often used interchangeably in the literature. RI is often used in academic scholarship whereas RRI is often used in European policy circles. In this article, I will use RI to denote the more general discussion of Responsible Innovation, and RRI to denote it's specific uptake by the European Commission.

[3] I will focus on the European RRI project as the main example of an active RI framework project.

step I suggest how RI can be further politicized through embracing a political conception of responsibility, questioning in turn whether responsible innovation is currently possible.

THE CRITICAL LANDSCAPE OF RI TODAY

Research and innovation processes are largely incentivized by economic interests. Critically reflecting on "what *should* steer innovation processes?", and, further, "what direction should innovation processes be steered in?" are precisely the kinds of normative questions that leading RI proponent Renè von Schomberg encourages. In "Why responsible innovation?" (2019), von Schomberg, critically reflects on the state of RI today. He highlights the major limitations that hinder RI from attaining what it sets out to do, including the task of steering innovation processes in the direction of *societally desirable outcomes*. At the same time, he advocates for – what I take to be – a stronger conception of RI by responding to these limitations with new visions. In this section, I will outline some of the main arguments of von Schomberg's work in order to set the critical landscape that RI finds itself in today.

The rapid development of RI is premised on the assumption that research and (scientific and technological) innovation practices lack, on their own, the incentive to take societal needs and desires into account. This assumption resulted from the widespread recognition in the 20th century that new and emerging technologies may have unpredictable and irreversible consequences that may be highly undesirable for both nature and society. As a result, institutional efforts were made to bridge the gap between scientific, societal, and ethical concerns by creating more concrete parameters for innovation.

Von Schomberg explains how the development and implementation of nuclear power plants in the 20th century served as a major catalyst in the efforts to create more concrete parameters for innovation:

> Nuclear power plants were regularly erected during the 1950s and 1960s with very little interference from our democratic institutions…. This occurred in the absence of professional risk governance and management, and in a culture of technological optimism. It was not only until the early 1970s, decades after the introduction of civil nuclear technology, that it was acknowledged that there were no solutions for the storage of nuclear waste. The institutionalization of risk identification and analysis as a distinct professional activity… emerged only at the end of the 1960s. (von Schomberg, 2019, p. 12)

This modern institutionalization of risk identification and analysis are incorporated in RI frameworks such as the European Commission's framework project of RRI. However, the current global implementation of RI frameworks is insufficiently guiding *responsible* innovation processes. Yet, what does it mean to say an innovation process

is responsible? According to von Schomberg an innovation process is responsible if, alongside risk identification and safety management, it is immediately directed toward attaining *societally desirable outcomes* (von Schomberg, 2019) – e.g., directly addressing pressing societal challenges such as environmental sustainability, health, or other welfare concerns. Furthermore, he argues that the 'right' direction of RI should be grounded in the normative anchor points found in the European Constitution and reflected in the European Framework Programme for Research and Innovation (von Schomberg, 2019, p. 16). At the same time, the 'right' direction of RI should be achieved in an ethical, sustainable, socially desirable, and democratic way (Owen *et al.*, 2012, p. 754).

The Deficits Hindering Strong RI Today

What then is hindering this form of *Responsible* Innovation? According to von Schomberg, global RI frameworks are facing several major deficits that need to be acknowledged and accommodated in order to develop stronger RI. One such deficit is the gap between government regulations and the market. This is because "[t]he state takes responsibility for the risks of products derived from new technologies, while the benefits are delegated to the market and defined in terms of success within the market" (von Schomberg, 2019, p. 14). Governments are primarily concerned with avoiding adverse effects of new and emerging technologies and insufficiently participate in steering innovations in a societally desirable and beneficial direction. In this sense, governments are engaging in a specific and narrow form of responsibility which leads to the market having increased power in deciding *what* is innovated and developed. The success of an innovation is thus largely expressed in terms of its profitability rather than an achievement of social, ethical, and political responsibilities towards citizens. As von Schomberg argues:

> Whereas public debates on the societal desirability of outcomes do not have a specific entry point in governmental policy-making, specific economic considerations drive the public and private funding of research and innovation actions. A first departure point for a vision of responsible innovation is therefore to advance governance mechanisms that could drive innovation to societally desirable ends. That is, instead of an exclusive focus on the risks of new technologies, the question of directing or redirecting research and innovation towards societally desirable ends has to be given importance in research and innovation programs. This implies that we not only need to have professional bodies for risk assessment but also professional bodies that should look into the type of outcomes we want to obtain from research and innovation processes, and the establishment of governance mechanisms that should give some direction to – or steer – the innovation process. (von Schomberg, 2019, p. 14)[4]

[4] Interestingly, while R. von Schomberg does in fact acknowledge that public debates – involved in determining the societal desirability of outcomes – do not have a specific entry point in the process of policy-making, he unfortunately does not develop this point. Given its political importance, this problem underlies and motivates my criticism of RI frameworks in this paper.

Thus, alongside the existing bodies concerned with risk and safety assessment, R. von Schomberg suggests (1) creating a professional body that explores and qualifies the desired *outcomes* of research and innovation processes and (2) establishing government mechanisms that would point innovation processes in the "right" direction (von Schomberg, 2019, p. 14).

The strong influence of economic incentives in deciding *what* gets innovated stands as an immediate hindrance to the development of RI. While there are countless new technologies and innovations entering the market, many of them lack actual societal significance (von Schomberg, 2019, p. 15). This is related to the commercial paradigm innovation processes find themselves in. Innovations and technologies are profitable, which from an economic viewpoint is a more directly *desirable* goal than the goal of "societally desirable outcomes". In this sense, economic interests function as a powerful incentive that influences various aspects of our existence.[5]

Economic incentives will point innovation processes towards achieving profit gains. On their own, however, these economic incentives will not ensure that innovations are steered in the "right" direction, e.g., *directly* benefitting humans, wildlife, or the environment more broadly speaking. Innovations that may have the direct *intention* of achieving societally desirable outcomes, but are deemed insufficiently profitable, are often sidelined completely or left for governments or philanthropic enterprises for further support (Owen & Pansera, 2019, p. 35; von Schomberg, 2019). In light of this, von Schomberg suggests that RI should strive to bridge the gap between the market and innovation processes intended to achieve societally desirable outcomes by, for instance, creating "new governance roles for public bodies and stakeholders" (von Schomberg, 2019, p. 15). Thus, instead of governments merely looking at the possible *consequences* of innovation (e.g., risk and safety consequences), they should be incentivized to include the standpoints and concerns of both the public and other relevant stakeholders.

R. von Schomberg further problematizes the macroeconomic model that is operative in the European Union which promotes scientific and technological advancements as ends in themselves. According to this model, it is not so much about *what* is being innovated, but rather *that* things are continuously being innovated. R. von Schomberg criticizes the European Union's lack of political initiative to hold innovation processes up to the normative standards that guide other European

[5] This last point is of course not a new insight but is part of a broader criticism of the incompatibility of capitalism and democratic politics which, although dating back to the 19th century, has only become more pronounced. This paper contributes to the ongoing effort to give politics a sufficient place amid a society largely governed by economic interests and powers.

policies. In light of this, he argues that RI should require "justification of the purpose and direction of innovation in terms of broadly shared public values" (von Schomberg, 2019, p. 17).

To briefly summarize, the current weaknesses of operative modes of R(R)I are largely due to its entanglement with global structures of profit incentivization. This entanglement results in RI insufficiently stimulating the development of innovations that are meant to directly address the pressing challenges of our time. In other words, public and private funding schemes currently do not ensure the development of *responsible* innovations. Visions for stronger RI therefore need to include (a) a recognition and accommodation of the deficits created by the structures of profit incentivization, (b) the *political* will and initiative to identify and enact core public values in innovation processes, and (c) management of the entire innovation process. To enable this form of RI, R. von Schomberg suggests, for instance, adding various professional bodies to the already existing governance mechanisms in order to address the current deficits that are hindering responsible innovation. These professional bodies, alongside other societal actors or stakeholders would democratically establish the *moral* evaluative criteria used to distinguish responsible and irresponsible innovation. Thus, for an innovation process to be deemed responsible, it would need to move towards these democratically achieved ethical standards. The suggested vision for RI that R. von Schomberg suggests would be there precisely to steer innovation in this (ethically) "right" direction.

The Invisibility of Politics in RI

While R. von Schomberg's diagnosis of the problems currently facing RI is accurate, the political dimension of the suggested visions for RI remains underdeveloped (Cf. Frodeman, 2019). RI has been criticized before for not being political *enough*. Michiel van Oudheusden presents a convincing criticism of RI frameworks claiming that they "largely ignore questions about the *politics in deliberation*…, as well as the *politics of deliberation*" (van Oudheusden, 2014, p. 68). The main concern regarding the politics involved *in* deliberation questions how consensus can be achieved in practice despite various power mechanisms at play (van Oudheusden, 2014, p. 73). This opens the discussion about the difficulties inherent in the deliberative process in general and further questions "how deliberation can be made sensitive to power dynamics and discursive exclusion that are facets of its constitutive and situated nature" (Owen & Pansera, 2019, p. 40). Regarding the politics *of* deliberation, van Oudheusden highlights how "the mere act of positing a common good reflects a politically motivated choice" (van Oudheusden, 2014, p. 73). Citing Igor Mayer, van Oudheusden asserts that a deliberative context is *always already* part of a particular history and worldview (van Oudheusden, 2014, p. 73). Specially in the context of RI, the ideas,

values, and concerns are pre-set. According to van Oudheusden, this preconceived and narrow conception of politics functions as an exclusionary mechanism in RI and therefore does not facilitate fair democratic practices. Van Oudheusden argues:

> …it would appear that participants who do not endorse deliberation or a commitment to deliberation and do not prioritize social and ethical concerns… over economic ones are placed on asymmetrical footing even before deliberation has officially begun. (van Oudheusden, 2014, p. 74)

For this reason, van Oudheusden asserts that this problematic aspect of the kind of politics present in RI needs to be acknowledged.

There is, however, still a lot of work to be done. I contend that, alongside RI's operative and problematic notion of politics, its conception of responsibility is also too narrow. In order to facilitate a broader and stronger conception of RI it is important to develop and incorporate a strong *political* conception of responsibility. This will enable RI to avoid the shortcomings of the primarily ethical conception of responsibility currently operative in RI. In what follows I contribute to the recent call for research to further develop an understanding of what the *political* dimension of *Responsible* Innovation could look like (Cf. van Oudheusden, 2014; Owen & Pansera, 2019; Reijers, 2020).

POLITICIZING RI: A DIALOGUE WITH HANNAH ARENDT

Attempts to define the concept of responsibility in the innovation context have been a longstanding challenge for researchers engaged in RI literature. In the last decade, much of the literature on RI has sought to move away from a consequentialist approach to responsibility and instead suggested a concept of responsibility that views innovation as a collective, uncertain, and unpredictable activity (Owen *et al.*, 2012, p. 756). Given that this strand of the literature strongly advocates for the principles of RI to be anchored in deliberative democracy, the conception of responsibility it puts forth is *value* – rather than rule-based (Owen *et al.*, 2012, p. 756), i.e., it strives to capture public *values* rather than set arbitrary normative standards. While there are certainly merits to a collective conception of responsibility, it risks neglecting the complex relationship between individual persons (citizens) and the world they experience and are inextricably a part of. This is important because the way in which persons *relate* to the world (e.g., through their work, their social status, religion, and so forth) also influences their experiences and critiques of it. Thus, to gain a deeper understanding of societal dissatisfactions and critique, European institutions should play a larger role in recognizing the relations and circumstances

that *motivate* critique.[6] Critically reflecting on these complex relationships can aid a preliminary understanding of *why* some persons may have different opinions about the same world, e.g., different preferences about what constitutes "societally desirable outcomes". Critically reflecting on this relationship, could better enable institutions such as the European Commission and its RI framework project to accommodate the dissatisfactions and concerns of its citizens. Engaging with socio-political and institutional critique in this way can further support the European Commissions' claims of implementing responsibility frameworks in the context of innovation that are democratically legitimate.

Alongside appreciating what motivates socio-political and institutional critique, it remains important to appreciate the politics *in* deliberation and the politics *of* deliberation. Heightened sensitivity toward the prevalence and influence of power mechanisms and existing ideologies in deliberation processes, can lead to a stronger framework of RI. Here too the role of motivation is important: e.g., considering what incentives persons or stakeholders have for pushing certain policies through or setting limited regulatory norms. The importance of this in the context of RI should not be underestimated. Consider for instance how some corporations or institutions may be motivated to innovate certain products over others due to possible profit or political gains. Take R. von Schomberg's example of how innovators in the medical field may be more motivated to create a treatment for a disease rather than a cure as there may be more financial gains to be made in the former case than the latter (von Schomberg, 2019, p. 15). Developing a treatment rather than a cure, due to the unprofitability of the latter, is not only morally but also *politically* questionable. Furthermore, since it is not in the best interest of society, such decisions can be said to be "societally undesirable".[7]

As it currently stands, innovation processes are largely motivated by possible profit gains. RI frameworks have been insufficiently incentivized to try to disentangle innovation processes from the economic paradigm they find themselves in. For that, political will and action is necessary. In other words, RI needs to be politicized. To explore how we can conceptualize this, I now turn to the work of political theorist and

---

[6] Robert Gianni argues that "we need to link the necessary moral responsibility to concrete social institutions in order to overcome the problems arising from a pluralist society and considering the necessity of promoting concrete and tangible measures." (Gianni, 2019, p. 64) I think Gianni makes a good point here, but I would extend this beyond moral responsibility. There should be a link between moral and political responsibility and European institutions. As Gianni further explains, "[r]esponsible efforts or practices cannot be put in place if they are not supported by specific institutional conditions, such as incentives, or rules framing the scope of research and innovation." (Gianni, 2019, p. 64) Here too the political dimension is important, alongside the moral one. It is therefore worth considering the extent to which the desired norms meant to "responsibly" guide innovation processes should be politically achieved.

[7] As we shall see, however, even the concept of "societally desirable outcomes" can be problematized for not being political enough, precisely because of the plurality inherent in society.

philosopher Hannah Arendt. We start with Arendt not only because her work has been very influential in the development of modern political theory, but also because she has dealt specifically with the question of responsibility.[8] What is perhaps most interesting however is the conception of political responsibility we can begin to develop *from* her work. Her work will thus provide a theoretical basis upon which we will discuss the importance of a political conception of responsibility in the RI context as well as contribute to recent efforts to explore and transform her theories and conceptions in relation to the phenomenon of RI (cf. Reijers, 2020). The main objective here is to plant the seeds for a stronger conception of RI through suggesting a *political* conception of responsibility. It is only by politicizing RI, I argue, that it can *respond* to the depoliticization[9] of the broader European public and in doing so further the critical reflection on how to legitimately obtain "societally desirable outcomes".

Hannah Arendt: The World of the Political

The development of new and emerging technologies has radically challenged operative Western liberal values, such as those of freedom and responsibility.[10] This has provoked renewed interest in Arendt's work and specifically her concepts of earth- and world-alienation (Cf. Berkowitz, 2018; Dinan, 2017). This is relevant here because the consequences of world-alienation in particular, as Arendt's conceives it, is a form of depoliticization. Given the scientific, socio-political, and philosophical developments since the publication of *The Human Condition* in 1958 (where Arendt deals with these concepts explicitly) Arendt's concepts should be rethought to suit our contemporary predicament. Furthermore, given Arendt's essayistic approach, interpretations of her concept of world are often rarely commented upon and

[8] Arendt's conception of responsibility is initially often linked to her coverage of the Eichmann trial and her criticism of those who operate within bureaucratic systems uncritically. Arendt's coverage of the Eichmann trial highlights the ease with which individuals can hide behind a bureaucratic curtain and thereby abstain from any form of moral or political responsibility. As Arendt reports, Eichmann famously argued that "[h]e did his duty, […]; he not only obeyed orders, he also obeyed the law" (Arendt, 2006 [1963], p. 135). Through arguing in this way, Eichmann abstained from taking any personal responsibility for his actions or their consequences. In this sense, Eichmann absolves himself of any guilt by diminishing his individual role in the greater system (portraying himself as 'just' a cog in a machine). While Arendt's analysis of Eichmann's intentions has been contested (Stangneth, 2011), the philosophical and political significance of her analysis still stands. Eichmann exemplifies a particular form of depoliticization precisely because he was unable "to think, namely, to think from the standpoint of somebody else" (Arendt, 2006 [1963], p. 49). In other words, he exemplifies the danger inherent in losing touch with the reality of our plural existence.

[9] Depoliticization is a broad term that can be used to denote "a decline in democratic, political creativity" (Straume and Humphrey, 2010, p. 10). In the work of Arendt, depoliticization arises when the world is seen from one dominant perspective, which covers over other possibilities of the world. As she argues, "the end of the common world has come when it is seen only under one aspect and is permitted to present itself in only one perspective" (Arendt, 1998 [1958], p. 58). It should be noted that when I use the term "depoliticization" here I presuppose that public political participation is not always a matter of personal choice. In other words, participation in the public realm is not always a possibility but is rather context dependent.

[10] Consider for instance Shoshana Zuboff's concern about the development of digital technology in The Age of Surveillance Capitalism (2019). There she argues that our notion of freedom has been compromised as a consequence of the commodification of our personal data.

therefore require clarification if they are to be applied to the RI context. Nonetheless, Arendt's work provides vital insights that, when re-interpreted in a critical and phenomenological context, are relevant today.[11] It is my contention that the RI discourse could benefit from taking this theoretical groundwork into account. Although Arendt does not explicitly provide practical guidelines for responsibility frameworks, her work can inspire us to critically reflect on the importance of including a political dimension of responsibility in European institutional frameworks such as RRI.

Since Arendt's concepts are idiosyncratic, we need to sufficiently clarify how we understand them from the outset. For instance, concepts integral to Arendt's work, such as "world" and "plurality", have several layers of meaning and need to be thoroughly unpacked in order to appreciate their role in her action-based political theory. In *Phenomenology of Plurality*, Sophie Loidolt systematically identifies and outlines Arendt's three-fold distinction of world. Loidolt distinguishes between (1) the 'appearing world', (2) the 'first in-between', and (3) 'the second in-between' (Loidolt, 2018, p. 98-99). What this all means will be elaborated here as these concepts of 'world' open the doors not only to Arendt's political theory, but also to a new development of a political conception of responsibility.

The Appearing World

The 'appearing world' refers to the most basic, phenomenological description of world in which "Being and Appearing coincide" (Arendt, 1978, p. 19; Cf. Loidolt, 2018, p. 98). Arendt argues that things *appear* by virtue of their existence – "[n]othing could appear, the word 'appearance' would make no sense, if recipients of appearances did not exist" (Arendt, 1978, p. 19). Here, appearance refers to a form of being "seen", but this includes all means of sense perception – i.e., sight, sound, taste, touch, smell. The world is thus fundamentally characterized by its active appearing-quality and directed

[11]Arendt's theory will be considered from a phenomenological viewpoint – whereby the unique interaction between the person and the world, and the political significance thereof, is taken into account. Phenomenological interpretations of Arendt's work are gaining momentum since Arendt's self-imposed distanciation from philosophy. In an interview with Gu¨nter Gaus in 1964, Arendt makes the famous claim: "I do not belong to the circle of philosophy. My profession, if one can speak of it at all, is political theory" (Arendt, 1994 [1964], p. 1). This statement has been quite influential in the reception of Arendt's thought, which has generally been taken up in "explicitly political terms" (Loidolt, 2018, p. 4). As a result, the strong philosophical dimension of her work -and specifically the (existential) phenomenological aspects – have been either neglected or completely ignored. There are of course important and notable exceptions. Dana Villa, for instance, published an influential book called Arendt and Heidegger: The Fate of the Political (Villa, 1996) which explores the strong Heideggerian themes of her work. However, Villa's work excludes important aspects of phenomenological thought, e.g., Husserl's work on empathy and intersubjectivity in Ideas II – which I deem important to understanding the connection between e.g., "the world of appearances", "disclosure", and "plurality", as Arendt conceptualizes them. Recently, scholars have been appreciating that Arendt is much more than just Heidegger's student and lover. As a result, works such as Sophie Loidolt's Phenomenology of Plurality (Loidolt, 2018) further pave the specifically phenomenological terrain upon which Arendt's work can be rethought.

toward someone that responds by *perceiving* it. As Loidolt remarks, Arendt's two other formulations of world – the 'first in-between' and the 'second in-between'[12]– are fundamentally anchored in this basic notion of appearance. The specific human activities that correspond to these formulations of world – i.e., work and action respectively –are very important, as it is precisely through them that appearance can gain its specific *meaningful reality*.

The First In-between: The World of Objects and Objectivity

The 'first in-between' refers to the tangible "world of objects and objectivity" (Arendt, 1998 [1958], p. 137; Loidolt, 2018, p. 98) that simultaneously relates and separates individuals from other individuals and the *objects* of their shared world. This 'first in-between' is artificial in that it is made by humans (in the mode of *homo faber*) and further conceptualized by them. This world is temporal and historical, kept in existence through continuous making/fabrication and remembrance.

Through *homo faber's* activity (i.e., the human activity of work), a social and material world is built. It is therefore on this level of human activity that institutions exist. Arendt argues that while humans create and build the world through the general practice of work, it also *conditions* them. While the natural Earth exists independently of human existence, the existence of the built-world depends *entirely* on humans; further, the built-world not only influences humans, but becomes a part of their existence. As Arendt writes:

> In addition to the conditions under which life has been given to man on earth, and partly out of them, men constantly create their own, self-made conditions which... possess the same conditioning power as natural things. Whatever touches or enters into a sustained relationship with human life immediately assumes the character of a condition of human existence. This is why men, no matter what they do, are always conditioned beings. Whatever enters the human world of its own accord or is drawn into it by human effort becomes part of the human condition. The impact of the world's reality upon human existence is felt and received as a conditioning force. The objectivity of the world – its object- or thing- character – and the human condition supplement each other; because human existence is conditioned existence, it would be impossible without things, and things would be a heap of unrelated articles, a non-world, if they were not the conditioners of human existence. (Arendt, 1998 [1958], p. 9)

In other words, the specific objects and practices of the human world shape and influence individuals, communities, and the status of human existence itself. The objectivity and practices of the world therefore create a sense of *meaningfulness*. This built-world relies on reification – which can also be understood as capturing ideas, stories, and events by materializing them in different ways. Once materialized, it is

[12] These two conceptions of world, corresponding to the activity of work and action respectively, are formulated most clearly in The Human Condition (1998 [1958]).

through remembrance that the world receives its specific historical dimension. The world therefore houses not only material objects but also makes possible, for instance, social practices, ideologies, cultures, and institutions. Through this world that is constantly being created, built, and remembered, individuals can relate to one another through their practical dealings with it. This allows individuals to refer and talk *about* the shared built-world, giving it its specific reality.

The built-world has a specific structuring dimension as well, simultaneously relating and separating individuals. For instance, when dealing with the built-world, individuals are on the one hand, concerned with the same *appearing-phenomenon*, but on the other hand, by virtue of the phenomenon standing *in-between* individuals, the specific way in which *it-seems- to-me* (Arendt, 2004, p. 433; Arendt, 1978, p. 21) remains unique.

> Living things *make their appearance* like actors on a stage set for them. The stage is common to all who are alive, but it *seems* different to each species, different to each individual specimen. Seeming – the it-seems-to-me, *dokei moi* – is the mode, perhaps the only possible one, in which an appearing world is acknowledged and perceived. To appear always means to seem to others, and this seeming varies according to the standpoint and the perspective of the spectators. [...] Seeming corresponds to the fact that every appearance, its identity notwithstanding, is perceived by a plurality of spectators. (Arendt, 1978, p. 21)

In other words, while we hold the world in common, we retain our unique perspective on it – recognizing that we are simultaneously equal and radically distinct from the other. According to Arendt, "only man can express this distinction and only he can communicate himself and not merely something"(Arendt, 1998 [1958], p. 176). In other words, despite dealing with the *same* object, individuals can recognize that they not only occupy a different objective-spatial perspective, but also a different subjective perspective.

The Second In-between: The World of the Political

Arendt's conceptualizations of speech and action are central to understanding her conception of the "second, subjective in-between" (Arendt, 1998 [1958], p. 183). According to Arendt, when persons speak and act – in a way that is novel and hence not just forms of idle talk or repetitions of "clichés"[13] – they disclose *who* they are. This disclosure is a political phenomenon that gains its specific reality through being seen and felt by others – as it actualizes a second in-between, occurring directly *between persons* (Arendt, 1998 [1958], p. 182). The specific reality actualized by speech and

[13] Political forms of speech and action are done for their own sake and are distinct from what Arendt calls "idle talk" (Arendt, 1998 [1958], p. 208). Here we can see similarities between Arendt concepts and those of Martin Heidegger found in Being and Time, including for instance: Rede (Cf., Arendt's "speech"), Gerede (Cf., Arendt's "idle talk"), and das Man (Cf., Arendt's conception of mass society or bureaucracy).

action is what I will refer to here as a *political* reality.[14] The 'second in-between' becomes the political space in which the person is immediately seen in their living reality before their words and actions are reified into familiar structures that *homo faber* can recognize. In other words, speech and action create an intersubjective in-between, a fleeting and intangible world that discloses unique persons, i.e., irreducible perspectives on the world. As Arendt explains:

> In acting and speaking, men show who they are, reveal actively their unique personal identities and thus make their appearance in the human world, while their physical identities appear without any activity of their own in the unique shape of the body and sound of the voice. (Arendt, 1998 [1958], p. 179)

In the act of speaking or acting individuals recognize one another as something *more* than the object-body. This something more is the *person*, in the political sense, or what Arendt famously calls the "who" someone is.[15] According to Arendt, "who" a person is cannot be captured in everyday language because words refer to something already *familiar and known*, while the "who" is always *unique and unfamiliar* (Arendt, 1998 [1958], p. 181). The expression of a person's uniqueness *is* a political gesture for Arendt. It requires courage as the outcome of action is risky given that the "who" that shows itself, is both unpredictable and irreversible. This unpredictability in turn is anchored in human plurality. The *appearance* of the who presupposes the other, an audience (as we know from the basic conception of world, "the appearing world"). Following from this:

> The disclosure of the 'who' through speech, and the setting of a new beginning though action, always falls into an already existing web [of human relationships] where their immediate consequences can be felt. … It is because of this already existing web of human relationships, with its innumerable, conflicting wills and intentions, that action almost never achieves its purpose[.] (Arendt, 1998 [1958], p. 184)

However, speech and action also need a *space* in order to appear to others. This space is what Arendt conceptualizes as a "space of appearances" (Arendt, 1998 [1958], p. 199), which is a political space par excellence. The political space of appearances constitutes a political reality that emerges when embodied individuals speak, act, and judge together. The emergence of this political space is thus the actualization of plurality's (political) potential. According to Arendt, "[w]hatever occurs in this space of

[14] This political reality appears between persons when they freely speak and act. It should be noted that for Arendt, freedom and politics are two sides of the same coin. She even argues that "[t]he meaning of politics is freedom" (Arendt, 2005 [1993], p. 108).

[15] I characterize it here as a recognition because in the disclosure, I recognize that the person, like myself, also has a unique and irreducible perspective on the world – their own it-seems-to-me.

appearances is political by definition, even when it is not a direct product of action" (Arendt, 1977 [1954], p. 155).

When a space of appearance does emerge and the temporal-historical "web" of human relationships becomes manifest, it can only remain so through *power*.[16] Put simply, the reality of the political space of appearances is dependent on persons coming together *through* action, yet this reality disappears as soon as individuals cease to be politically engaged (whereby the political space of appearances is dismantled). In a telling passage Arendt writes:

> Power preserves the public realm and the space of appearances, and as such it is also the lifeblood of the human artifice, which, unless it is a scene of action and speech, of the web of human affairs and relationships and the stories engendered by them, lacks its ultimate *raison d'être*. Without being talked about by men and without housing them, the world would not be a human artifice but a heap of unrelated things to which each isolated individual was at liberty to add one more object; without the human artifice to house them, human affairs would be floating, as futile and vain, as the wandering nomad tribes. (Arendt, 1998 [1958], p. 204)

In other words, the space of appearances and the public realm are connected. The former provides the latter with its source of inspiration. This inspiration can then materialize or be reified into, for instance, institutions and social practices. While the human world can persist without the political, that world would become futile, static, and without novel change. One form of depoliticization occurs when the world is seen from one dominant viewpoint whereby political action does not appear as a possibility and a political space of appearance cannot emerge.[17] This is the form of depoliticization that Arendt devotes much of her work addressing. While the phenomenon of depoliticization is complex and certainly not limited to Arendt's conceptualization of it, her understanding of it highlights a problem(s) that occurs when economic incentives primarily motivate actions – and in this context, innovation processes.

Let us briefly summarize the discussion thus far. Following Loidolt's distinction, Arendt has a three-fold conception of world. At the most basic level, the world *appears* and thus presupposes a sentient being to whom it appears. The basic forms of appearance and plurality that typify the 'appearing world' anchor the first in-between (the world of objects and objectivity; the built-world) and the second in-between (the world of the political). The first in-between is created by humans and

[16] Arendt has an idiosyncratic understanding of power. She argues that "[p]ower is actualized only where word and deed have not parted company, where words are not empty and deeds not brutal, where words are not used to veil intentions but to disclose realities, and deeds are not used to violate and destroy but to establish relations and create new realities." (Arendt, 1998 [1958], p. 200) Power is thus a dynamic potential that keeps the spirit of action in existence.

[17] For more on the phenomenon of "depoliticization" and links to the work of Arendt, see Straume and Humphrey, 2010.

provides structure, stability, and familiarity to human existence – simultaneously conditioning it. The second in-between breaks with the familiarity created and safeguarded by *homo faber*[18]. Speech and action, which actualize the second in-between, result in the introduction of something new by means of *showing* a previously unseen perspective or position on the world – i.e., it shows something novel, challenging the familiar ways in which the world is understood. By virtue of plurality, this speech and action is thrown into a temporal-historical world composed of "innumerable, conflicting wills and intentions". This can provoke new dimensions to the meaning of words or concepts, institutions, or systems. Although a person's irreducible subjective experience of the world can never actually be inhabited/embodied by another person, political speech and action articulate that experience and make it accessible.

POLITICIZING RI THROUGH A POLITICAL CONCEPTION OF RESPONSIBILITY

As we saw in section one, RI is facing several challenges including (1) the impact of economic incentives in steering innovation processes and (2) its political shortcomings. It is my contention that the conception of responsibility in the RI context should include a stronger political dimension. Specifically in the context of the European Commission, and framework projects such as RRI, we find an operative conception of politics that is too narrow. By including and considering different "stakeholders", the European Commission takes itself to be sufficiently political in its democratic processes – and by extension politically responsible. To challenge this, I presented an interpretation of Arendt's conception of the political to facilitate a theoretical insight into the complex dynamics of plurality. From the interpretation presented, we can understand the political as an actualized state of plurality (Loidolt, 2019), set into motion by speech and action. At the same time, the political space of appearances – which results from such an actualization – is contingent and relies on the continuous political participation of persons. These steps have been important to be able to start our reflections on why, and in which ways, RI is insufficiently political. In this following section, we will look at how our analysis thus far can aid us in politicizing RI though a political conception of responsibility.

[18] Here I refer for instance to homo faber's attempts to understanding something as something or a person in terms of what they are rather than who they are. It should further be noted that Arendt's conceptions of world are inter-related. For instance, while Arendt's conceptions of speech and action have an inherent character of novelty and spontaneity, they still spring from the human-built world (i.e., the world of homo faber). Further, it is often the world of homo faber that provides the subject matter and sources of inspiration for action.

It has been suggested before that Arendt's political theory *entails* a particular conception of responsibility. The link has been made, for instance, between Arendt's theory of disclosure (i.e., the disclosure of 'who' someone is) and an implicit concept of responsibility (Williams, 2015); or, as Loidolt argues, that there is an implicit ethics "that springs from *actualizing plurality*" (Loidolt, 2019, p. 234). Arendt's theory does indeed provide us with a wealth of tools to develop a conception of responsibility rooted in a conception of the political. In my opinion, what makes Arendt's theory so unique is the way in which it accounts for the person (i.e., the irreducible 'who') without forsaking the idea of something shared (i.e., plurality).[19] Here it is important to recall that for Arendt action almost never achieves its aim due to the dynamics of actualized human plurality. This has implications for a conception of responsibility, namely, that it cannot end with the individual (political) act. Rather, it encapsulates the reciprocity that takes place between persons and the (plural) world.

Following these reflections on Arendt's theory of the political, political responsibility can be understood *as the enactment and maintenance of a political space of appearances*. In our specific context, if we understand political responsibility in this way, then we must urge the European Commission to critically reflect on the democratic character of European political processes. The reflection on this process should take into consideration the dissatisfaction of *all* inhabitants of European Member States – across *all* social and political classes. This would include recognizing the motivations and constraints that encourage or discourage persons to participate in the political process. If the European Commission critically reflects upon the democratic character of its political processes and actively recognizes these motivations and constraints, then it would have to acknowledge the *real* phenomenon of depoliticization and the discontent that undermines the efforts of its own framework projects.

Yet, this form of political responsibility is often limited in European institutions and governments. Political responsibility is often neglected in favor of economic responsibility. The European Commission's RRI framework is caught in this problematic as well; its conception of responsibility remains too narrow – often primarily focused on economic and moral forms of responsibility. As we saw earlier, the market largely decides what gets innovated, and by virtue of the market's impersonal nature it *cannot* question whether or not those innovations are positively impactful for society or the environment. While RI tries to accommodate this gap –

[19] One possible criticism to Arendt's political theory – or more specifically, to her conception of action – is that it insufficiently considers the material conditions necessary for persons to engage in politics. Given the scope of this paper, we will not be able to address this limitation here. However, this is arguably a weakness of Arendt's work and deserves to be further developed.

between the market and "societally desirable" impacts – it avoids altering the very structure of the market-driven society it operates in. Even if, as R. von Schomberg suggests, RRI guides innovation processes according to the pre-established European values inscribed in the European Constitution (and consolidated more recently in the Lund Declaration), it still evades its political responsibility as I have outlined it here. In the RRI framework project, the political occurs within specific parameters, i.e., in a controlled, institutional environment. Van Oudheusden argues that if RI presents itself "to policy-makers as a politically neutral tool [it] risks trivializing and undermining the very policy changes RI advocates seek to instigate" (van Oudheusden, 2014, p. 81). He further argues that RI could, and perhaps should, adopt "a more politically laden language of agendas, interests, impacts, and power" (van Oudheusden, 2014, p. 81). Building on this, I argue that the concerns of the world's citizens, and their real-life experiences of the world (which motivate political speech and action), should be actively recognized, if we are to speak of political responsibility. Policy frameworks of the European Commission, such as the RRI framework, are often experienced as emanating from an ivory tower – detached from the real-life experiences of "ordinary citizens". By becoming more sensitive to the importance of political responsibility, RI can become more inclusive by responding to the actual needs and concerns of citizens. My suggestion therefore encourages RI to further reflect on its political shortcomings in order to properly distinguish responsible from irresponsible innovation processes and credibly argue for "societally desirable outcomes".

CONCLUSION

In this paper I have addressed some of the major difficulties RI is facing with regard to, on the one hand, the power of profit motives to steer innovation processes and on the other hand, the lack of a strong political dimension in its operative conception of responsibility. Alongside these difficulties, RI struggles to address the depoliticization of persons and societies which highlights the lack of a strong political dimension in its conception of responsibility. In light of this, I argued that RI needs to be further politicized. In this endeavor to broaden the political scope of RI, I have drawn on Arendt's political theory to shed light on how responsibility and politics can be understood as two sides of the same coin. RI needs to address widespread societal depoliticization by adopting a conception of responsibility that is sufficiently *political* in nature; only then can it claim to steer innovations in directions desired *by* society and more specifically, everyday persons. While there is still a great deal of work to be done, further critical reflection is urgently needed on both the political dimensions of responsibility and the impacts of depoliticization on the very possibility of *responsible* innovation.

REFERENCES

Arendt, H. (2006 [1963]). *Eichmann in Jerusalem: A Report on the Banality of Evil*. Intro. by Amos Elon. NY: Penguin Books.

Arendt, H. (2004). Philosophy and Politics. *Social Research*, 71(3), 427-454.

Arendt, H. (1998 [1958]). *The Human Condition*. Ed. by Margaret Canovan. 2nd Ed. Chicago, IL: The University of Chicago Press.

Arendt, H. (1978). *The Life of the Mind*. San Diego: Harcourt, Inc.

Arendt, H. (2005). *The Promise of Politics*. Ed. by Jerome Kohn. NY: Schocken Books.

Arendt, H. (1977 [1954]). What is Freedom? In *Between Past and Future: Eight Exercises in Political Thought* (pp. 143-171). Dallas, PA: Penguin Books.

Arendt, H. (1994). 'What Remains? The Language Remains': A Conversation with Günter Gaus. In *Essays in Understanding 1930-1954: Formation, Exile, and Totalitarianism* (pp.1-23). Ed. by Jerome Kohn. NY: Schocken Books.

Berkowitz, R. (2018). The Singularity and the Human Condition. *Philosophy Today*, 62(2), 337-355.

Blok, V., & Lemmens, P. (2015). The Emerging Concept of Responsible Innovation. Three Reasons Why it is Questionable and Calls for a Radical Transformation of the Concept of Innovation. In B. Koops, J. van den Hoven, H. Romijn, T. Swierstra, & I. Oosterlaken (Eds.) *Responsible Innovation 2: Concepts, Approaches and Applications* (pp. 19-36). Dordrecht: Springer.

Dinan, M. (2017). The Robot Condition: Karel Čapek's R.U.R. and Hannah Arendt on Labor, Technology, and Innovation. *Perspectives on Political Science*, 46(2), 108-117.

European Commission (2014). *Responsible Research and Innovation: Europe's Ability to Respond to Social Challenges*. Directorate-General for Research and Innovation,

Publications Office. https://op.europa.eu/en/publication-detail/-/publication/2be36f74-b490-409e-bb60-12fd438100fe

Fleming, R. C., & Mauger, R. (2021). Green and Just? An Update on the 'European Green Deal'. *Journal for European Environmental & Planning Law*, 18(1-2), 164-80. https://doi.org/10.1163/18760104-18010010

Frodeman, R. (2019) Review of International Handbook on Responsible Innovation. A Global Resource. *Journal of Responsible Innovation*, 6(2), 255-257.

Gianni, R. (2019). Choosing freedom: ethical governance for responsible research and innovation. In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation: A Global Resource* (pp. 49-69). UK: Edward Elgar Publishing Limited.

Godin, B. (2015). *Innovation Contested*: *The Idea of Innovation over the Centuries*. Routledge.

Grunwald, A. (2018). Diverging Pathways to Overcoming the Environmental Crisis: A Critique of Eco- Modernism from a Technology Assessment Perspective. *Journal of Cleaner Production*, 197(2), 1854-1862.

Heidegger, M. (2010 [1927]). *Being and Time*. Trans. by Joan Stambaugh. Albany: State University of New York Press.

Loidolt, S. (2018). *Phenomenology of Plurality: Hannah Arendt on Political Intersubjectivity*. NY: Routledge.

Owen, R., & Pansera, M. (2019). Responsible innovation: process and politics. In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation: A Global Resource* (pp.35-48). UK: Edward Elgar Publishing Limited.

Owen, R, Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy*, 39(6), 751-760.

Pianta, M., & Lucchese, M. (2020). Rethinking the European Green Deal: An Industrial Policy for a Just Transition in Europe. *Review of Radical Political Economics*, 52(4), 633-641.

Stangneth, B. (2016). *Eichmann before Jerusalem: The Unexamined Life of a Mass Murderer*. Repr. London: Vintage.

Stilgoe, J,. Owen, R., & Macnaghten, P. (2013). Developing a Framework for Responsible Innovation. *Research Policy*, 42(9), 1568-1580.

Straume, I. S., & Humphrey, J. F. (2010). Introduction. In I. S. Straume & J. F. Humphrey *Depoliticization: The Political Imaginary of Global Capitalism* (pp. 9-19). NSU Press.

van Oudheusden, M. (2014). Where are the politics in responsible innovation? European governance, technology assessments, and beyond. *Journal of Responsible Innovation*, 1(1), 67-86. https://doi.org/10.1080/23299460.2014.882097

Villa, D. (1996) *Arendt and Heidegger: The Fate of the Political*. Princeton, NJ: Princeton University Press.

von Schomberg, L., & Blok, V. (2019). Technology in the Age of Innovation: Responsible Innovation as a New Subdomain Within the Philosophy of Technology. *Philosophy & Technology*, 34, 309-323. https://doi.org/10.1007/s13347-019-00386-3

von Schomberg, R. (2019). Why responsible innovation? In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation: A Global Resource* (pp.12-32). UK: Edward Elgar Publishing Limited.

Reijers, W. (2020). Responsible innovation between virtue and governance: revisiting Arendt's notion of work as action. *Journal of Responsible Innovation*, 7(3), 471-489.

Williams, G. (2015). Disclosure and Responsibility in Arendt's *The Human Condition*. *European Journal of Political Theory*, 14 (1), 37-54.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: PublicAffairs.

Anticipatory responsible innovation. Futures construction in the face of the techno-economic imperative[1]

Hannot Rodríguez*, Sergio Urueña and Andoni Ibarra*****

* *University of the Basque Country*

** *University of the Basque Country*

*** *University of the Basque Country*

ABSTRACT

The call for the development of more responsible research and innovation has increasingly permeated European Union research and development policies. Specifically, under the auspices of approaches such as "Responsible Research and Innovation" (RRI) and "Open Science", these policies conceive of the need to make innovation dynamics radically open and debatable, even with regard to the underlying preferences and expectations shaping them. Responsibility has thus been conceived in eminently anticipatory terms, that is, in terms of collectively taking care in the present of the futures enabled through innovation practices. This normative conception, which emphasises the politicisation of the ways futures are constructed through innovation and goals they are oriented towards, is nonetheless realised within a context where the prevailing way of approaching the future with regard to innovation systems is highly committed to a capitalist imperative of technological progress and economic growth. This article argues that while anticipation – understood as an interventive practice – can deploy valuable responsibilisation heuristics, their degree of disruptiveness, or openness, may depend on how such interventive practice engaging with futures deals with this techno-economic commitment, or imperative.

Keywords: Anticipation; Responsible Innovation; RRI; Open Science; Anticipatory Ambivalence; Socio-Technical Futures.

Proposal Submitted 27 August 2020, Article Received 9 April 2021, Reviews Delivered 5 July 2021, Revised 13 August 2021, Accepted 5 Setember 2021, Available online 28 February 2022.

---

[1] Acknowledgement: This work was supported by the project PID2020-114279RB-I00, funded by the Spanish Ministry of Science and Innovation and the Spanish State Research Agency, MCIN/AEI/10.13039/501100011033. It was also supported by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FFI2015-69792-R), the Vice-rectorate for Research of the University of the Basque Country UPV/EHU (PPGA19/23, and GIU19/051), and the Basque Government's Department of Education (IT1205-19). Sergio Urueña acknowledges the support of a FPI predoctoral fellowship (BES-2016-079192), funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund. The authors would very much like to thank the two reviewers. Their insightful comments and suggestions proved to be of great help to improve and clarify this article.

INTRODUCTION

Research and innovation policies have shown a growing interest over the past two decades in promoting "responsible innovation" dynamics that transcend mere expert assessment and evaluation of risks and "collateral impacts" associated with – often seen as autonomous – scientific and technological progress (von Schomberg & Hankins, 2019). In this regard, for example, "Horizon 2020", the 8th European Union (EU) Framework Programme for Research and Innovation (2014-2020), claimed to address responsibility according to a Responsible Research and Innovation (RRI) approach, whereby innovation processes are conceived as susceptible to being radically open and debatable, even with regard to the preferences and expectations underlying them (European Commission, 2013). Similarly, the following, more recent 9th EU Framework Programme, "Horizon Europe" (2021-2027), claims to adhere to an "Open Innovation, Open Science, Open to the World" perspective, arguably committed to radical knowledge sharing and promoting robust science-society coproduction commitments and dynamics in the era of globalised information and communication technologies (European Commission, 2019a).

This *a priori* commitment to opening-up the actual dimension of innovation processes and their goals – including ways and conditions to achieve them – to public scrutiny would open the door to the possibility of problematising the mechanisms that deliberately narrow and reify the socio-technical futures considered desirable and plausible, and whereby current scientific-technological practices are legitimised and modulated (Jasanoff & Kim, 2015). In this respect, the issue of more inclusive governance of socio-technical futures seems to have become a characteristic attribute of Responsible Innovation (Stilgoe *et al.*, 2013, p. 1570) and other well-established related approaches such as Technology Assessment (Grunwald, 2019), which aim to democratise research and innovation.

The future is thus considered a key element within innovation dynamics. This implies that any serious attempt to democratise research and innovation practices and trajectories must allow for the problematisation and broadening of the set of socio-technical futures at stake. In this respect, it has been argued that anticipation, broadly understood as a practice characterised by the use of the future to orient present actions, functions as an interventive resource to democratise future representations that colonise and orient the present (Arnaldi, 2018; Stemerding *et al.*, 2019; Yoshizawa, 2019). Here, for instance, anticipation aims to engender alternative practices for action, challenge the *status quo* and enhance emancipation (Withycombe *et al.*, 2019). According to Guston (2008, p. vi), Anticipatory Governance is "a broad-based capacity extending through society that can help individuals and

institutions act on a variety of inputs to manage emerging knowledge-based technologies while such management is still possible".[2]

However, this article argues that while anticipation – understood as an interventive socio-epistemic practice – can deploy valuable heuristics to responsibilise innovation, the degree of disruptiveness, or openness, of such heuristics would be severely limited by the prevailing manner of approaching the future in the context of innovation systems such as the EU. These systems are very much dominated by a techno-economic imperative. This means that research and innovation's predominant (and *de facto* almost indisputable) mission is to achieve certain prefixed industrial and economic milestones. Such an imperative expresses a strong commitment to a technocratic and economicist vision of technological progress, and aligns with a techno-capitalist ideological approach (Beckert, 2016; Godin, 2016; Shelley-Egan *et al.*, 2020). Anticipation's disruptive potential will thus depend on how it is conceived and used under such conditions. In this sense, it proves crucial to analyse whether – and how – anticipatory practices enable futures envisioning capable of critically scrutinising the techno-economic imperative's normative base and building alternative relations between innovation and economic dynamics.

The scope and meaning of anticipation must therefore always be analysed and elucidated in relation to the specific and situated socio-political contexts where anticipatory practices take place. It is according to these contexts, and the ways in which anticipation operates within them, that anticipation can tend to act either as a *disruptive* tool – i.e., at the disposal of the critical-reflexive openness of socio-technical systems – or, on the contrary, as a *limiting* element – i.e., focused on orienting science and technology governance towards normative milestones that are prefixed and impervious to debate. In this respect, this article identifies and characterises this ambivalent feature of anticipation in relation to its potentially dual, "disruptive-limiting" role in the context of innovation systems such as the EU's. The economistic imperatives underlying EU research and innovation policies seem to hinder the development of more disruptive, or open, anticipatory practices, which are characteristic of more radically inclusive interpretations of proposals such as RRI or Open Science (Gerber *et al.*, 2020).

To this end, this article is structured as follows: first, the emergence and meaning of RRI and Open Science in the context of EU innovation system is explained.

---

[2] The four pillars of Anticipatory Governance are: foresight, engagement, integration, and "ensemble-isation" (or coordinated mobilisation of the three previous pillars). Anticipation, or critical engagement with the future, is explicitly operationalised through foresight, which "aims to enrich futures-in-the-making by encouraging and developing reflexivity in the system" (Barben *et al.*, 2008, p. 986).

It argues that there is a dominant tendency within this system to make certain interests and values prevail (as well as certain assumptions about their desirability and feasibility) in relation to innovation and its dynamics, which seems to limit the inclusive and transformative potential of such proposals. On the basis of this analysis, the relevance of anticipation in scientific-technological modulation processes is then discussed, including the possibility of conceiving and articulating anticipatory governance mechanisms intentionally aimed at fostering the critical-collective construction of future representations and promoting alternative courses of action accordingly, *in the present*. These considerations go hand in hand with recognition of the aforementioned situated and necessarily ambivalent character of anticipation, and of the difficulties associated with attempts to promote eminently disruptive, or inclusive, anticipatory dynamics in the context of innovation systems. They are systems that are highly committed to – and constrained by – the techno-economic imperative linked to the ideology of techno-industrial developmentalism. Finally, the main conclusions are presented.

RESPONSIBLE INNOVATION IN TIMES OF RRI AND OPEN SCIENCE

The EU innovation system has radicalised its narratives on "responsible innovation" over the last two decades to the extent that its most recent formulations conceive it in terms of the degree of inclusiveness, or integration, of a heterogeneity of actors and publics (Eizagirre *et al.*, 2017; Macnaghten, 2020).

Thus, the 8th Framework Programme for Research and Innovation, "Horizon 2020" (2014-2020), via its RRI approach linked to the "Science with and for Society" (SwafS) initiative, set its deliberative intention and capacity as the main characteristic of a responsible innovation process, implying that even the values, motivations and expected benefits of innovations should be subject to public scrutiny (European Commission, 2013). This would involve transcending the dominant institutional tendency to impose regulatory frameworks on technological innovations whose social justification is unproblematised (Felt *et al.*, 2007; Owen *et al.*, 2013). According to the European Commission (EC), RRI "allows all societal actors (…) to work together during the whole research and innovation process" (European Commission, 2013, p. 4).

The most recent 9th EU Framework Programme, "Horizon Europe" (2021-2027), also recognises the need to promote "better linkages between scientists, citizens and policy-makers" (European Commission, 2018a, p. 74). "Horizon Europe" is in fact conceived as a means of promoting a radically open, or participatory and transparent, innovation system characterised by "the three Os": "Open Innovation, Open Science and Open to the World". This "open" initiative aspires to facilitate free access to

knowledge and knowledge sharing, "where new knowledge is created through global collaborations involving thousands of people from across the world and from all walks of life" (Moedas, 2015, p. 1) in order to achieve scientific excellence and innovative efficiency (Bogers *et al.*, 2018; European Commission, 2016, 2019a).

However, it seems pertinent to question the meaning and transformative scope of this type of initiatives in view of the risk of their instrumentalisation by an innovation system whose ultimate, or main, objective appears to be the industrial exploitation of knowledge (Godin, 2016; Shelley-Egan, Gjefsen & Nydal, 2020). In other words, it is worth questioning the motivations and types of imperatives guiding such initiatives (Fiorino, 1989). Their transformative, or "opening-up", potentials, should not therefore be reified, nor taken for granted. Rather, light should be shed on how the framings, power instantiations and instrumentalisation dynamics pervading innovation systems tend to foreclose, or "close down", the emergence of alternative ways of appraising and executing technological progress (Stirling, 2008).

Significantly, RRI has also been characterised according to a limited set of dimensions (namely: public engagement, gender equality, open access to research, science education, and ethics) that are subordinate to achieving the goals of "mak[ing] science more attractive (…), raise the appetite of society for innovation, and open up further research and innovation activities" (European Commission, 2013, p. 4). Under this "practical" characterisation (i.e., RRI "[i]n practice" [European Commission, 2013, p. 4]), public participation, for example, has been conceived as being aimed at "reinforcing public confidence in science" (European Parliament and Council of the EU, 2013, p. 106).[3] Similarly, it is argued that "[t]he European Union will not remain competitive at the global level unless it promotes Open Science, and relatedly, Open Innovation" (European Commission, 2018b, p. 4), which might also be indicative of a risk of excessive instrumentalisation of the "open" ideal (Mayer, 2015), whereby citizens are mainly represented as actors with "a central and transversal role to play in bringing innovation to the market" (European Commission, 2016, p. 17).[4]

These proposals in favour of more radically inclusive, or "open", responsible innovation must therefore be measured in light of the fundamental tension between

[3] The "Horizon 2020" website dedicated to Public Engagement in Responsible Research and Innovation expresses itself in similar non-disruptive, or "normalising", terms. See: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/public-engagement-responsible-research-and-innovation (consulted on 11 February 2021).

[4] Citizen science is presented as one of the pillars of Open Science (European Commission, 2018b; Mendez *et al.*, 2020), as an activity that "aims to encourage the inclusion of non-institutional participants, in other words the general public, in the scientific process" (European Commission, 2016, p. 53), and that aspires to "re-direct research agendas towards issues of concern to citizens" (European Commission, 2016, p. 54). At the same time, however, "Citizen Science is often linked with outreach activities, science education or various forms of public engagement with science as a way to promote Responsible Research and Innovation" (European Commission, 2016, p. 54).

the demands for more socially responsible techno-industrial progress and the political and epistemic dynamics that are firmly committed to developmentalism and competitiveness (Owen & Pansera, 2019; Rodríguez *et al.*, 2019; Stirling, 2016; von Schomberg & Blok, 2021).

Arguably, this tension is made even more evident and aggravated by the dominant institutional tendency to minimise its scope by assuming that techno-industrial progress is compatible with a broad set of socio-environmental considerations. This is in line with the use of a prefixed normative framework (namely: "Promotion of scientific and technological advance", "Competitive social market economy", "Promotion of social justice", "Sustainable development", "Quality of life, high level of protection" [von Schomberg, 2013, p. 58]),[5] and the possibility of its harmonisation, which seems to be taken for granted. This relates to the fact that research within "Horizon 2020" and "Horizon Europe" is not organised according to disciplinary criteria but – as recommended by the Lund Declaration (2009) – according to "a challenge-based approach" (Council of the EU, 2013, p. 966), where "[r]esearch and innovation are key drivers of sustainable growth and industrial competitiveness, and they will contribute to finding solutions to today's problems" (European Commission, 2018a, p. 17). Such challenges, in all their heterogeneity (i.e., economic, social, environmental and health-related), could thus all be solved (via science and technology leadership) together in a constitutively compatible way (e.g., ERA Expert Group, 2008, p. 36). In EC terms:

> "(…) Horizon Europe will strengthen the Union's scientific and technological bases in order to help tackle the major global challenges of our time and contribute to achieving the Sustainable Development Goals (SDGs). At the same time, the programme will boost the Union's competitiveness, including that of its industries. (…) Europe's success increasingly depends on its ability to transform excellent scientific results into innovation that have a real beneficial impact on our economy and quality of life (…)" (European Commission, 2018c, p. 1).

"Having it all at once" therefore seems possible here. It is considered that the unwavering commitment of innovation systems to the techno-economic imperative – whereby technological development is conceived as a key element for economic growth and competitiveness, and therefore an absolute priority – is compatible with other interests and concerns (moreover, it would follow that incompatibility in its strict sense is not even an option). This techno-economic imperative arguably frames the series of interests and concerns appraised inside innovation systems (Godin, 2016; Shelley-Egan *et al.*, 2020).

[5] This prefixed normative framework emanated from the 1992 EU Treaty, or "Maastricht Treaty" (von Schomberg, 2013, p. 56-58).

Therefore, this ideal of equilibrium, or harmony, would not conflict with initiatives such as RRI or Open Science, with all their attributed disruptive potential. Rather, the ideal would be reinforced, given that "RRI fosters the creativity and innovativeness of European societies" (European Commission, 2013, p. 4), "open science will (…) increase the innovation potential of results generated by Union funding" (European Commission, 2018c, p. 14) and, ultimately, "(…) when partners from across academia, industry, public authorities and citizen groups are invited to participate in the research and innovation process [according to an open science policy], creativity and trust in science increases"[6].

In this respect, it is worth considering the need to analyse the way in which – albeit seemingly paradoxical at first sight – a political approach where responsibility is related to inclusivity and heterogeneity restricts, or limits, the capacities and will to develop alternative ways of innovating and relating to science and technology – precisely because it goes hand in hand with a normative horizon whose content and compatibility are prefixed. Thus, rather than posing a "postmodern" or "ideologised" threat to scientific autonomy and the authority of expert knowledge (e.g., Kuntz, 2012, 2017), such inclusivist initiatives seem to serve a certain triumphalist, or radically enlightened, image of science and technology: they are entrusted with the mission of leading the resolution of major socio-environmental challenges by assuming, as a matter of principle, the capacity to satisfy a heterogeneous set of values and challenges without incurring traumatic renunciations, as noted above.

THE FUTURE, AND SCIENCE AND TECHNOLOGY GOVERNANCE

The actions and programmes mobilised in the EU around initiatives such as RRI and Open Science are implemented via a set of public policies that are often vulnerable to tensions, as seen. This stands in the way of achieving the objectives linked to more responsible research and innovation processes, i.e., processes that are more open to the consideration of a plurality of issues, interests and criteria (Novitzky *et al.*, 2020). Coordinated action to promote inclusive research and innovation practices must thus take into consideration and value such difficulties under whose terms these practices acquire a constitutively *ambivalent* character. This is because they can be interpreted and function both as enabling, or "disruptive", resources, and as elements subordinate to a set of imperatives and assumptions that constrain their potential to bring about substantive changes in the trajectories of techno-industrial progress.

---

[6] European Commission: "The EU's open science policy"; available at: https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science_en (consulted on 29 January 2021).

In this context, the issue of purposes and motivations proves fundamental. That is, the desirability of inclusively responsible science and innovation *per se* is not the only consideration to be taken into account; some link must be established with for what *purpose* such openness is encouraged, i.e., with the *rationale* for making science and innovation more open. In this regard, and, in particular, the question of objectives, and their associated representations of the future, proves critical when analysing the meaning and scope of this trend towards openness (Jasanoff & Kim, 2015). Given that "Horizon 2020" and "Horizon Europe" approach research to solve societal challenges and demands, it is appropriate to consider how these challenges and demands are determined, as well as the proposals to address them.

This requires attention to be focused on the socio-technical futures construction and establishment processes that progressively guide (responsible) science and technology governance. Such futures are generally presented by the institutional domain as highly promising in socio-economic terms and, in this sense, function as orientational, legitimising and promotional elements of their associated innovation dynamics (Jasanoff & Kim, 2015; Schiølin, 2020). This, of course, does not imply that the futures are impervious to controversy surrounding both their alleged benefits and their potential health, environmental and social risks (Jasanoff, 2016).[7] Future representations and the normative assumptions accompanying them are thus a constitutive element of scientific-technological development, just as they are a constitutive element of modern societies (which are marked by a clear rationalistic-calculative impetus) where such developments unfold (Giddens, 1990; Hölscher, 1999).

The performative character of socio-technical futures in the present is explained in light of the phenomenon of anticipation. Anticipatory action (both individual and social) is considered by Anticipation Studies to be any action performed – whether consciously or unconsciously – on the basis of a representation, or model, of the future (Poli, 2017; Poli & Valerio, 2019). According to this definition, every socio-technical system can be considered an anticipatory system, insofar as it is co-inhabited by a series of future representations that influence the network of heterogeneous actions which jointly and progressively constitute its co-production and co-evolution (Konrad *et al.*, 2016; Lösch *et al.*, 2019).

Konrad *et al.* (2016) show that two cases of anticipatory practices can be distinguished. On the one hand, some anticipations occur *de facto* in socio-technical

[7] The case of strategic emerging technologies is particularly significant in this respect. This is an area where tensions between highly enabling scientific-technological novelty and concerns and reticence about the futures associated with this novel (and uncertain) potential are particularly pronounced (Alvial-Palavicino & Konrad, 2019; Rodríguez, 2018).

systems. Here anticipations represent the constellation of actions and decisions that, informed by future representations such as visions (Lösch *et al.*, 2019) imaginaries (Jasanoff & Kim, 2015) and expectations (Alvial-Palavicino & Konrad, 2019; Borup *et al.*, 2006), help shape research and innovation commitments and trajectories. On the other, certain other anticipations take on an explicitly normative-instrumental character. In this second case of anticipatory practice, the future is intentionally and interventively mobilised (according to explicitly designed methodological criteria and aims) in order to promote more responsible innovation (Arnaldi, 2018; Selin, 2011). According to this latter instrumental-interventive version, anticipation becomes an instrument, or resource, that serves to modulate (somewhat) "more responsible" innovation. However, different paradigms, or normative models, of responsibility co-exist around what is considered "being responsible", configuring different functional modes of engagement with the future and, therefore, different modes of operationalising anticipation (i.e., of "using the future") (Adam & Groves, 2007).

The mainstream mode of anticipatorily governing innovation systems has technocratic flavours, which consist of articulating actions defining governance on the basis of expert-based models of the future, whose objective is to project the system's future state as accurately as possible. This anticipatory activity can be carried out for various purposes, such as changing the course of events to prevent the predictive model from being fulfilled, or developing adaptive strategies that seek to reduce or accelerate forecasted impacts.

Understood under this predictivist model, anticipation has been considered within innovation systems as a particularly useful instrument to promote the development of these systems that is sensitive to their potential impacts and consequences. A fairly clear example of this is risk analysis of techno-industrial developments, which seeks to assess and manage *ex ante* both socio-technical accidents (e.g., Perrow, 1984), and progressive and cumulative health and ecological impacts (e.g., Cranor, 2017). Thus, risk analysis points directly to the evils of progress – allowing, or legitimising, its critique and regulation – (Delogu, 2016) while at the same time denoting, under its institutionalized form, the – disputable – assumption that the risks of techno-industrial progress can be foreseen and regulated (i.e., controlled) without having to forgo economic growth and consumerism (Dickson, 1984, p. 261-306).

There are, however, alternative ways of using anticipation. Thus, within certain approaches that are committed to developing more inclusive and responsible research and innovation dynamics, anticipatory activity does not seek to be based on models of the future with predictive pretensions, and explicitly aims to distinguish itself from such models (Stilgoe *et al.*, 2013, p. 1571; Barben *et al.*, 2008, p. 985). Many

of the scholars and practitioners, who now – more or less tentatively (Kuhlmann *et al.*, 2019; Fisher, 2019) – promote the use of anticipation as part of their respective innovation governance models, recognise the complexity of carrying out predictivist claims (e.g., Guston & Sarewitz, 2002). They also point to the counterproductive nature of these predictive claims, given their tendency to shield the normative assumptions that, *de facto*, underlie them (e.g., Sarewitz *et al.*, 2000). Anticipation is, instead, conceived of here as a practice aimed at the collective problematisation of future states deemed *(im)plausible* and *(un)desirable* (Guston, 2014; Selin, 2011) in order to generate a series of capability-building heuristics that enable a more reflexive intervention in the present (Konrad *et al.*, 2016, p. 479-483; Ramos *et al.,* 2019; Rip, 2018, Chapter 2). Understood in this way, it is not surprising that anticipation is considered a defining dimension of Responsible Innovation. In fact, Responsible Innovation is defined in terms of "taking care of the future through collective stewardship of science and innovation in the present" (Stilgoe *et al*, 2013, p. 1570).

It should be noted, however, that this last characterisation of anticipation can acquire various degrees of radicality depending on the diversity of anchoring assumptions considered *prefixed* in practice (i.e., not susceptible to scrutiny, as seen in the following section). This disruptive version of anticipation is by no means alien to ambivalences either.

THE AMBIVALENT POTENTIAL OF ANTICIPATION, AND THE TECHNO-ECONOMIC IMPERATIVE

Anticipation can be used both to support the dynamics of radically inclusive RRI, or "open"-like responsible innovation and to promote more instrumentalised variants limited by certain predefined regulatory frameworks and practices (Ruggiu, 2019). In fact, the meanings and performative potentials and inclinations of responsible innovation normative initiatives are, in general, constitutively contextual. This means that the transformative capabilities of their defining principles, which include anticipation, can arguably be understood as being a function of the manners in which such principles are approached and used in accordance with different preferences, commitments and power relations. Responsible innovation frameworks and principles are therefore constitutively *ambivalent*. They can both help to "open up" research and innovation practices to a more plural set of perspectives and concerns and help to "close them down" on the basis of certain pervasive technocratically-oriented assumptions (e.g., "value-free" science, a sharp "expert/lay" epistemic divide, and instrumentalised public participation), and always according to particular "context and implementation" conditions (Stirling, 2008, p. 268).

Anticipation can thus act both as an "opening-up", or *disruptive*, element as well as a "closing-down", or *limiting*, element depending on how it deals with the prevailing situated framings and dynamics. Thus, the disruptive variant of anticipation would use the collective problematisation of a system's future states in order to facilitate the emergence of alternative courses of action, while the limiting version would assume the desirability and plausibility of certain future scenarios from the outset in order to proceed with the problematisation of potential impacts (both positive and negative) that could occur during the realisation of such scenarios. Therefore, in contrast to "disruptive anticipation", aimed at "opening-up" the range of action alternatives in the present, "limiting anticipation" focuses on exploring both the different consequences that could emanate from the realisation of a given future project, and the possible contingencies that may affect (e.g., impede, hinder or enhance) its achievement.

The ambivalent feature of the abovementioned "anticipatory heuristics" is merely a reflection of the tensioned nature of innovation systems in their dealings with responsible innovation and its demands. This ambivalence is ultimately evident in the divergent ways in which anticipation operates in the context of innovation systems that are highly committed to a set of techno-economic imperatives at the disposal of techno-industrial developmentalism, and its associated economic growth. There is therefore an urgent need to analyse the way in which this "disruption-limitation" ambivalence is expressed in relation to such imperatives. Ultimately, the disruptive or limiting degree of anticipatory activity is a function of the system's (in)capacity to develop anticipatory resources that enable the critical-reflexive re-elaboration of its normative foundations, which constrain the degree of openness of socio-technical alternatives.

Understood from a radically inclusive, or open, perspective, anticipation means, as mentioned, "opening-up" the discussion to the plurality of future projects that the various social actors might hold. This discussion is precisely the resource expected to feed the process of imagining alternatives for action (Lehoux *et al.*, 2020). The plurality of future visions and projects held by the heterogeneous social actors (i.e., the diversity of different knowledges, expectations, interests and normativities) is used as a heuristic resource to provide orientation and enrich the present (Grunwald, 2013). According to this disruptive conception, anticipation can be characterised as a practice directed towards the collective problematisation of future states.

In this vein, anticipatory knowledge here is not knowledge about the future *per se* (as pointed out above), i.e., "is not about seeing into the future (prudence) or saying what the future is going to be (prediction) or estimating the chances of a certain outcome (probabilistic forecasting)" (Foley *et al.*, 2018, p. 228). Thus, rather than

aspiring to mitigate and eliminate uncertainty about the future, disruptive anticipatory knowledge embraces (both empirical and normative) uncertainty. The future is not a space to be epistemically and technically conquered. Rather, it is a politically open, debatable reality. So, the value of anticipatory knowledge, now, lies in its ability to facilitate the identification of alternative ways of acting by producing heterogeneous representations of socio-technical futures. Producing such heterogeneous futures representations depends on the full consideration and interrelation of a diversity of different knowledges, values and political preferences. This implies that the robustness of anticipatory knowledge may arguably be understood as a function of the level of integration, or inclusiveness, of anticipatory processes. In these terms, anticipation would then be a genuine source of robust knowledge, namely, knowledge "able to withstand variety and interference", and produced through "interactions and struggles" (Rip, 2018, p. 21).

Based on this understanding of anticipatory knowledge, the epistemic legitimisation of futures representations does not rely on the probabilistic production and assessment of future scenarios. Rather, it relies on the less constraining demand of plausibility (Selin, 2011; Wiek *et al.*, 2013). Plausibility thus turns into a crucial epistemic device as it enables the envisioning and consideration of futures that would otherwise be excluded under a probability-based stance (Ramírez & Selin, 2014). Plausibility navigates between the probable and the possible; it is more inclusive than probability and more constrictive than possibility. Even though plausibility conceptually enables a broader and richer set of future scenarios to be considered, the extent and detail of such scenarios will depend on the variety of assumptions embraced (or not) when determining plausibility in practice. In this sense, the extent of variety and detail of scenarios is conditioned by the understandings of "the (im)plausible" and "the (un)desirable" (not) actually considered and mobilised during plausibility negotiation processes. Thus, the reception of plural conceptions concerning "the (im)plausible" and "the (un)desirable" is the epistemic device that enables futures to be opened up (Urueña, 2019).

Anticipation and negotiation of "the (im)plausible" and "the (un)desirable" will not, however, be free of resistances. Anticipation can only be an efficient and realistic tool if it takes into account that, from the very outset, governance processes and forms "are not without tensions" (Siune *et al.*, 2009, p. 4). This is due to the fact that the degree of (im)plausibility and (un)desirability is decided contextually, as pointed out above. Thus, the anticipatory heuristics' degree of radicality (in terms of "disruption-limitation") will, among other factors, depend heavily on the elements considered prefixed when alternatives are envisioned (i.e., elements delimiting *ex ante* the domain of futures under consideration) (Urueña, 2019). In this sense, questioning which assumptions constrain these anticipatory mechanisms within innovation governance

proves a crucial issue, as does how, by whom and why they are mobilised and established as such.

One normative element that frames, constrains and limits what is considered (im)plausible is the techno-economic imperative. This imperative characterises innovation systems such as the EU's, as seen. These systems are firmly based on the ideology of techno-industrial developmentalism and the constitutively related assumption of absolute (and, on principle, unquestionable) harmony, or compatibility, between the different concerns and interests analysed above. Any anticipatory futures-building process constrained by this imperative (whether consciously or unconsciously assumed) will reflect socio-technical assemblages where innovation systems are geared towards maximising economic growth. This ideological frame limits, or impoverishes, the future. Conceived as an achievable, or "designable", state of affairs, the future under industry-driven anticipatory thinking and practice is not approached as an opportunity to reflect on and debate alternative socio-technical scenarios and trajectories or new normative horizons. Instead, it serves the purpose of pre-legitimising certain techno-industrial priorities and projects, precisely under the assumption that the future itself is susceptible to being instrumentally (i.e., technically) mastered and controlled (Nordmann, 2010, 2014).

Anticipation can only become a disruptive practice if it is able to envision alternative futures where relations between innovation practices and the market are articulated through alternative values that reach beyond economic rationalisation. In other words, only anticipatory practices capable of challenging the prevailing plausibility frameworks that take the techno-economic imperative for granted will take on a radically disruptive character. The degree of success of anticipatory heuristic practices could in turn modulate the socio-technical arrangements, revealing different gradients of intensity (i.e., they could modulate the socio-technical arrangements according to different gradients of openness and closure). Although anticipation is seen as an emancipatory interventive instrument, it should be interpreted within the complex socio-technical network where it emerges and intended to prove effective. Anticipatory disruptive practices aim to be functional within networks where forces typically resistant to change exist, and where actors tend to perpetuate the status quo (Withycombe *et al.*, 2019). The broad capitalist anchoring and *momentum* of our societies will not only hinder the envisioning and emergence of alternative modes of relations between innovation and the market, but also the very conception of "uses of the future" that are not framed and pragmatically oriented towards increasing profit (Beckert, 2016).

In this sense, even the more disruptivist forms of anticipation (i.e., those working under the less constraining epistemic register of plausibility) are vulnerable

to ambivalences. Such anticipations may trigger the opening-up of the future with regard to certain aspects while, at the same time, they may close it down in relation to certain other aspects. For instance, problematising the potential future impacts associated with nanotechnologies may obscure the relevant debate on whether nanotechnologies are desirable in themselves. The occurrence of nanotechnology would be taken for granted here (i.e., in principle, it would be considered plausible), implying that any future scenario excluding nanomaterials would *de facto* be ignored. Thus, the apparently disruptive anticipatory exercise would in fact align with more committed pro-nanotechnology policy and industry narratives (which, in turn, are often guided by the techno-economic imperative).

It is therefore important to recognise that the transformative potential of anticipation as an instrument of intervention is significantly constrained by the broader framework of understanding and action within which it aims to become operational in order to foster more responsible innovation. In this sense, the European innovation system's commitment to anticipation, on the basis of which the EC argues, for example, that "[RRI] implies anticipating and assessing potential implications and societal expectations with regard to research and innovation" (European Commission, 2013, p. 4), seems to reflect the limiting-type version of anticipation rather than the disruptive variant. Thus, for example, the SwafS line of research "Developing Inclusive, Anticipatory Governance for Research & Innovation" claims to serve the development of "scenarios regarding possible future RRI activities and how these activities are perceived by science and society" (European Commission, 2017, p. 8-9) in order to "contribute to inclusive and anticipatory governance in the context of strategic priority-setting for future R&I (funding) policy in Europe" and help "the strengthening of the research and innovation ethics framework" (European Commission, 2017, p. 9), meaning that anticipation is subordinate to a mapping of the various actors' perception with respect to the RRI framework itself, and limited to a set of prefixed priorities that are impervious to criticism. Another SwafS research initiative, "Building the knowledge base for SwafS", on the other hand, addresses anticipation as a resource linked to the exercise of examining the ways science and society co-evolve. This includes analysis of potential social attitudes towards this very co-evolution, provided that "[u]nderstanding the co-evolution of science and society will help proactive and anticipatory policy making" (European Commission, 2019b, p. 43), so anticipatory activity seems to be identified here with a strategy for minimising socio-technical uncertainty.[8]

[8] This does not mean that some characterisation of anticipation with open and disruptive traits cannot be found within the EC. For instance, according to certain discourses within the EC, research and innovation should "play an ever-more important role in creating the future we want", by "opening the

Promoting radically more responsible innovation requires recognising this ambivalence of anticipation as a tool for modulating "responsible" practices and analysing different "uses of the future" and the rationale underlying them. At the same time, this ambivalence needs to be studied in a context where, despite the existence of various narratives seeking a more radical opening-up of innovation systems (articulated in more disruptive conceptions of anticipation), the disruptive potential of these narratives is limited in number and scope by innovation dynamics which are significantly compromised by the techno-economic imperative characterising modern capitalist societies.

CONCLUSION

Anticipation, when acting as an interventive resource aimed at enabling the problematisation and collective production of representations of socio-technical futures legitimising and guiding current scientific-technological practices, is able to function as a heuristic tool with the potential to promote more responsible research and innovation dynamics. This responsibility, in line with a certain type of more radical interpretation of recent proposals for "responsible innovation", such as RRI or Open Science, within the EU's innovation system framework, is defined in terms of inclusiveness and openness with respect to the interests, means and goals underlying research and innovation dynamics (i.e., with respect to the processes and elements determining what lines of scientific and technological action should be promoted and the milestones they should be oriented towards).

This article has sought to show, however, that this "responsible heuristic", serving to facilitate radically open and heterogeneous debate on the very purposes and interests underlying innovation systems, must necessarily be approached by taking into account that the anticipatory heuristic's degree of "disruptiveness" or "limitation" will depend on the way in which anticipation, as an interventive practice, plays out in the context of innovation systems such as the EU's. These systems, which are deeply committed to economic-industrial developmentalism and the subsequent instrumentalisation of science and technology dynamics, pursue the achievement of certain milestones related to economic growth and competitiveness.

This techno-economic imperative thus acts as an element that constrains and limits the envisioning of alternative futures; it represents, in other words, the system's (in)capacities to develop anticipatory resources that enable the critical-reflexive re-

discussion on future research and innovation policy and investment, and (…) promoting engagement and participation by society in the policy process" (Moedas, 2017, p. 7).

elaboration of its normative foundations. These foundations are based on the assumption of absolute (and, on principle, unquestionable) harmony, or compatibility, among different societal concerns and interests with regard to techno-industrial progress, its problems and the unquestionable and urgent *de facto* quest for economic development and profit. In this sense, it has been argued that there is therefore the need to analyse whether – and how – anticipatory practices enable futures envisioning capable of critically scrutinising the techno-economic imperative's normative base, and of building alternative relations between innovation and economic dynamics.

All of this implies that the scope and meaning of anticipatory practice must always be approached in relation to the specific socio-political contexts where this practice takes place. Depending on these contexts, anticipation will tend to act either as a *disruptive* instrument (i.e., serving critical-reflexive openness of socio-technical systems) or, on the contrary, as a *limiting* element (i.e., focused on orienting the governance of science and technology towards prefixed normative milestones that are impervious to debate).

This article has therefore proceeded to identify and characterise this ambivalent feature of anticipation in relation to its potentially dual, "disruptive-limiting" role in the context of EU innovation system. This is a context where the dominant, economicist imperatives severely hinder the possibility of developing more disruptivist anticipatory practices (i.e., practices in line with the most radically inclusive interpretations of "responsible innovation" proposals such as RRI or Open Science).

REFERENCES

Adam, B., & Groves, C. (2007). *Future Matters: Action, Knowledge, Ethics*. Brill.

Alvial-Palavicino, C., & Konrad, K. (2019). The rise of graphene expectations: Anticipatory practices in emergent nanotechnologies. *Futures*, 109, 192-202.

Arnaldi, S. (2018). Retooling Techno-Moral Scenarios. A Revisited Technique for Exploring Alternative Regimes of Responsibility for Human Enhancement. *NanoEthics*, 12(3), 283-300.

Barben, D., Fisher, E., Selin, C., & Guston, D. H. (2008). Anticipatory Governance of Nanotechnology: Foresight, Engagement, and Integration. In E. J. Hackett, O. Amsterdamska, M. Lynch & J. Wajcman (Eds.), *The Handbook of Science and Technology Studies. Third Edition* (pp. 979-1000). The MIT Press.

Beckert, J. (2016). *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Harvard University Press.

Bogers, M., Chesbrough, H., & Moedas, C. (2018). Open Innovation: Research, Practices, and Policies. *California Management Review*, 60(2), 5-16.

Borup, M., Brown, N., Konrad, K., & van Lente, H. (2006). The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*, 18(3-4), 285-298.

Council of the EU (2013). Council Decision of 3 December 2013 establishing the specific programme implementing Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020) and repealing Decisions 2006/971/EC, 2006/972/EC, 2006/973/EC, 2006/974/EC and 2006/975/EC (2013/743/EU). *Official Journal of the European Union* (20.12.2013), L 347, 965-1041.

Cranor, C. F. (2017). *Tragic Failures: How and Why We are Harmed by Toxic Chemicals*. Oxford University Press.

Delogu, B. (2016). *Risk Analysis and Governance in EU Policy Making and Regulation: An Introductory Guide*. Springer.

Dickson, D. (1984). *The New Politics of Science*. Pantheon Books.

European Commission (2013). *Horizon 2020, Work Programme 2014–2015: 16. Science with and for Society*, C(2013) 8631 of 10 December 2013. Brussels.

European Commission (2016). *Open Innovation, Open Science, Open to the World – a vision for Europe*. Publications Office of the European Union.

European Commission (2017). *Horizon 2020, Work Programme 2016-2017: 16. Science with and for Society*, C (2017)2468 of 17 April 2017. Brussels.

European Commission (2018a). *Annexes to the Proposal for a Decision of the European Parliament and of the Council on establishing the specific programme implementing Horizon Europe – the Framework Programme for Research and Innovation*, COM(2018) 436 final (7.6.2018). Brussels.

European Commission (2018b). *OSPP-REC: Open Science Policy Platform Recommendations*. Publications Office of the European Union.

European Commission (2018c). *Proposal for a Regulation of the European Parliament and of the Council establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination*, COM(2018) 435 final (7.6.2018). Brussels.

European Commission (2019a). *Orientations towards the first Strategic Plan implementing the research and innovation framework programme Horizon Europe*. Brussels.

European Commission (2019b). *Horizon 2020, Work Programme 2018-2020: 16. Science with and for Society*, C (2019)1849 of 18 March 2019. Brussels.

Eizagirre, A., Rodríguez, H., & Ibarra, A. (2017). Politicizing Responsible Innovation: Responsibility as Inclusive Governance. *International Journal of Innovation Studies*, 1(1), 20-36.

ERA [European Research Area] Expert Group (2008). *Challenging Europe's Research: Rationales for the European Research Area (ERA)* (EUR 23326). Office for Official Publications of the European Communities.

European Parliament and Council of the EU (2013). Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 – the Framework Programme for Research and Innovation (2014-2010) and repealing Decision No 1982/2006/EC. *Official Journal of the European Union* (20.12.2013), L 347, 104-173.

Felt, U. (Rapporteur), Wynne, B. (Chairman), Callon, M., Gonçalves, M. E., Jasanoff, S., Jepsen, M., Joly, P.-B., Konopasek, Z., May, S., Neubauer, C., Rip, A., Siune, K., Stirling, A., & Tallacchini, M. (2007). *Taking European Knowledge Society Seriously. Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate, Directorate-General for Research, European Commission* (EUR 22700). Office for Official Publications of the European Communities.

Fiorino, D. J. (1989). Environmental Risk and Democratic Process: A Critical Review. *Columbia Journal of Environmental Law*, 14(2), 501-547.

Fisher, E. (2019). Governing with ambivalence: The tentative origins of socio-technical integration. *Research Policy*, 48(5), 1138-1149.

Foley, R. W., Guston, D. H., & Sarewitz, D. (2018). Towards the anticipatory governance of geoengineering. In J. J. Blackstock & S. Low (Eds.), *Geoengineering our Climate? Ethics, Politics, and Governance* (pp. 223-243). Routledge.

Gerber, A., Forsberg, E.-M., Shelley-Egan, C., Arias, R., Daimer, S., Dalton, G., Cristóbal, A. B., Dreyer, M., Griessler, E., Lindner, R., Revuelta, G., Riccio, A., & Steinhaus, N. (2020). Joint declaration on mainstreaming RRI across Horizon Europe. *Journal of Responsible Innovation*, 7(3), 708-711.

Giddens, A. (1990). *The Consequences of Modernity*. Stanford University Press.

Godin, B. (2016). Making sense of innovation: from weapon to instrument to buzzword. *Quaderni*, 90, 21-40.

Grunwald, A. (2013). Modes of orientation provided by futures studies: making sense of diversity and divergence. *European Journal of Futures Research*, 2(1), 30.

Grunwald, A. (2019). The inherently democratic nature of technology assessment. *Science and Public Policy*, 46(5), 702-709.

Guston, D. H. (2008). Preface. In E. Fisher, C. Selin & J. M. Wetmore (Eds.), *The Yearbook of Nanotechnology in Society. Volume 1: Presenting Futures* (pp. v-viii). Springer.

Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social Studies of Science*, 44(2), 218-242.

Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society,* 24(1-2), 93-109.

Hölscher, L. (1999). *Die Entdeckung der Zukunft*. Fischer Taschenbuch Verlag.

Jasanoff, S. (2016). *The Ethics of Invention: Technology and the Human Future*. W. W. Norton & Company.

Jasanoff, S., & Kim, S.-H. (Eds.) (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. The University of Chicago Press.

Konrad, K., van Lente, H., Groves, C., & Selin, C. (2016). Performing and Governing the Future in Science and Technology. In U. Felt, R. Fouché, C. A. Miller & L. Smith-Doerr (Eds.), *The Handbook of Science and Technology Studies.* Fourth Edition (pp. 465-493). The MIT Press.

Kuhlmann, S., Stegmaier, P., & Konrad, K. (2019). The tentative governance of emerging science and technology—A conceptual introduction. *Research Policy*, 48(5), 1091-1097.

Kuntz, M. (2012). The postmodern assault on science. *EMBO reports*, 13(10), 885-889.

Kuntz, M. (2017). Science and Postmodernism: From Right-Thinking to Soft-Despotism. *Trends in Biotechnology*, 35(4), 283-285.

Lehoux, P., Miller, F. A., & Williams-Jones, B. (2020). Anticipatory governance and moral imagination: Methodological insights from a scenario-based public deliberation study. *Technological Forecasting & Social Change*, 151, 119800.

Lösch, A., Grunwald, A., Meister, M., & Schulz-Schaeffer, I. (Eds.) (2019). *Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges*. Springer VS.

Lösch, A., Heil, R., & Schneider, C. (2019). Visionary Practices Shaping Power Constellations. In A. Lösch, A. Grunwald, M. Meister & I. Schulz-Schaeffer (Eds.), *Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges* (pp. 67-88). Springer VS.

Lund Declaration (2009). Europe Must Focus on the Grand Challenges of our Time. In *New Worlds – New Solutions. Research and Innovation as a Basis for Developing Europe in a Global Context* (The Swedish EU Presidency Conference – Final report), Uno Svedin (Rapporteur), Lund (Sweden), 7-8 July 2009, pp. 40-41.

Macnaghten, P. (2020). *The Making of Responsible Innovation*. Cambridge University Press.

Mayer, K. (2015). From Science 2.0 to Open Science - Turning rhetoric into action? *STCSN E-Letter*, 3(1). http://stcsn.ieee.net/e-letter/stcsn-e-letter-vol-3-no-1/from-science-2-0-to-open-science

Mendez, E. (Chair), Lawrence, R. (Editor and Coordinator), MacCallum, C. J. (Writing Chair), Moar, E. (Rapporteur) *et al.* (2020). *Progress on Open Science: Towards a Shared Research Knowledge System. Final Report of the Open Science Policy Platform*. Publications Office of the European Union.

Moedas, C. (2015). Open Innovation, Open Science, Open to the World (SPEECH/15/5243). *"A new start for Europe: Opening up to an ERA of Innovation" Conference*. 22 June 2015. Brussels.

Moedas, C. (2017). Foreword. In A. Ricci, C. Sessa, & M. Weber (Authors), *New Horizons: Future Scenarios for Research & Innovation Policies in Europe* (European Commission, Directorate-General for Research and Innovation) (p. 7). Publications Office of the European Union.

Nordmann, A. (2010). A forensics of wishing: technology assessment in the age of technoscience. *Poiesis & Praxis*, 7(1-2), 5-15.

Nordmann, A. (2014). Responsible innovation, the art and craft of anticipation. *Journal of Responsible Innovation*, 1(1), 87-98.

Novitzky, P., Bernstein. M. J., Blok, V., Braun, R., Chan, T. T., Lamers, W., Loeber, A., Meijer, I., Lindner, R., & Griessler, E. (2020). Improve alignment of research policy and societal values. *Science*, 369(6499), 39-41.

Owen, R., & Pansera, M. (2019). Responsible Innovation and Responsible Research and Innovation. In D. Simon, S. Kuhlmann, J. Stamm, & W. Canzler (Eds.), *Handbook on Science and Public Policy* (pp. 26-48). Edward Elgar.

Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A Framework for Responsible Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 27-50). Wiley.

Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. Basic Books.

Poli, R. (2017). *Introduction to Anticipation Studies*. Springer.

Poli, R., & Valerio, M. (2019). *Anticipation, Agency and Complexity*. Springer.

Ramírez, R., & Selin, C. (2014). Plausibility and probability in scenario planning. *Foresight*, 16(1), 54-74.

Ramos, J., Sweeney, J. A., Peach, K., & Smith, L. (2019). *Our futures: by the people, for the people*. Nesta.

Rip, A. (2018). *Futures of Science and Technology in Society*. Springer VS.

Rodríguez, H. (2018). Nanotechnology and Risk Governance in the European Union: the Constitution of Safety in Highly Promoted and Contested Innovation Areas. *NanoEthics*, 12(1), 5-26.

Rodríguez, H., Eizagirre, A., & Ibarra, A. (2019). Dynamics of responsible innovation constitution in European Union research policy: tensions, possibilities and constraints. In R. von Schomberg & J. Hankins (Eds.), *International Handbook on Responsible Innovation: A Global Resource* (pp. 167-180). Edward Elgar.

Ruggiu, D. (2019). Models of Anticipation Within the Responsible Research and Innovation Framework: the Two RRI Approaches and the Challenge of Human Rights. *NanoEthics*, 13(1), 53-78.

Sarewitz, D., Pielke, R. A., Jr., & Byerly, R. (2000). *Prediction: Science, Decision Making, and the Future of Nature*. Island Press.

Schiølin, K. (2020). Revolutionary dreams: Future essentialism and the sociotechnical imaginary of the fourth industrial revolution in Denmark. *Social Studies of Science*, 50(4), 542-566.

Selin, C. (2011). Negotiating Plausibility: Intervening in the Future of Nanotechnology. *Science and Engineering Ethics*, 17(4), 723-737.

Shelley-Egan, C., Gjefsen, M. D., & Nydal, R. (2020). Consolidating RRI and Open Science: understanding the potential for transformative change. *Life Sciences, Society and Policy*, 16, 7.

Siune, K. (Chair), Markus, E. (Rapporteur), Calloni, M., Felt, U., Gorski, A., Grunwald, A., Rip, A., de Semir, V., & Wyatt, S. (2009). *Challenging Futures of Science in Society. Emerging trends and cutting-edge issues*. Publications Office of the European Union.

Stemerding, D., Betten, W., Rerimassie, V., Robaey, Z., & Kupper, F. (2019). Future making and responsible governance of innovation in synthetic biology. *Futures*, 109, 213-226.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568-1580.

Stirling, A. (2008). "Opening Up" and "Closing Down". Power, Participation and Pluralism in the Social Appraisal of Technology. *Science, Technology, & Human Values*, 33(2), 262-294.

Stirling, A. (2016). Addressing scarcities in responsible innovation. *Journal of Responsible Innovation*, 3(3), 274-281.

Urueña, S. (2019). Understanding "plausibility": A relational approach to the anticipatory heuristics of future scenarios. *Futures*, 111, 15-25.

von Schomberg, L., & Blok, V. (2021). The turbulent age of innovation. *Synthese*, 198(Suppl. 19), 4667-4683. https://doi.org/10.1007/s11229-018-01950-8

von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 51-74). Wiley.

von Schomberg, R., & Hankins, J. (Eds.) (2019). *International Handbook on Responsible Innovation: A Global Resource*. Edward Elgar.

Wiek, A., Withycombe Keeler, L., Schweizer, V., & Lang, D. J. (2013). Plausibility indications in future scenarios. *International Journal of Foresight and Innovation Policy*, 9(2-4), 133-147.

Withycombe Keeler, L., Bernstein, M. J., & Selin, C. (2019). Intervening Through Futures for Sustainable Presents: Scenarios, Sustainability, and Responsible Research and Innovation. In A. Lösch, A. Grunwald, M. Meister & I. Schulz-Schaeffer (Eds.), *Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges* (pp. 255-282). Springer VS.

Yoshizawa, G. (2019). Reflexive Hermeneutics Against Closing Down Technology Assessment Discourses: The Case of Synthetic Biology. In A. Lösch, A. Grunwald, M. Meister & I. Schulz-Schaeffer (Eds.), *Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges* (pp. 189-210). Springer VS.

Open access, responsibility and the "platformization" of academic publishing[1]

Raúl Tabarés*

* *TECNALIA, Basque Research and Technology Alliance (BRTA)*

ABSTRACT

Digitalization was supposed to be a transformation force for the academic publishing sector, but it has reinforced the oligopoly of for-profit academic publishers. Open access (OA) was also meant to counterbalance this situation, but after a decade of efforts it seems that it has not achieved their goals. This essay explores how the combination of digitalization and OA have contributed to reinforce the lock-in effects exerted in the sector by digital platforms operated by for-profit academic publishers. I also explore alternative paths for the development of OA with the theoretical lenses that provide responsible innovation, putting social emphasis at the politics and values that lie at the heart of academia. I argue that exploitation, appropriation of labor and quantification metrics widely present in this social domain must be counterbalanced with different actions that do not focus alone in making freely available scientific articles for citizens.

Keywords: Digital Platforms; Responsible Innovation; Platform Economy; Platform Capitalism; Open Access; Open Science.

Proposal Submitted 28 August 2020, Article Received 17 April 2021, Reviews Delivered 2 July 2021, Revised 24 August 2021, Accepted 3 November 2021, Available online 28 February 2022.

[1] Acknowledgement: This work has been possible thanks to the funds received from the New HoRRIzon Horizon 2020 project under Grant Agreement number 741402 and PLUS Horizon 2020 project under Grant Agreement number 822638.

INTRODUCTION

Responsible Innovation (RI) has revealed itself during the last decade as a transformative force that can enhance, extend and strength science-society interactions (Owen & Pansera, 2019; Stilgoe *et al.*, 2013). Its normative vision, namely Responsible Research and Innovation (RRI), has been pushed forward by the European Commission (EC) in the last decade for driving innovation towards socially desirable ends (von Schomberg, 2013). Open access (OA) has been one of the "six keys" promoted by the EC in this normative vision (European Commission, 2012) to maximize the visibility and availability of scientific articles and assuring that citizens will not pay twice (first for conducting the research and second for reading its results) (Delaney *et al.*, 2020). OA was later also included in the open science (OS) policy of the EC[2] introduced at later stages of the Horizon 2020 Framework Programme for Research and Innovation, and as a step towards the adoption of RRI in the EU research ecosystem (European Commission, 2016).

However, and after several years of OA development, its full implementation in the academic publishing sector is still far away. At the same time, digitalization was also to be considered a transformation force for the industry, but it has not achieved the significant change that was thought. Lowering operation costs of journals, launching new OA journals and favoring OA implementation due to new possibilities that can confer digital technologies to publishing were among its initial aims, but this process has not provided the aforementioned results. The aim of this essay is to shed some light on these controversies and outlining the challenges that the combination of OA and the "platformization" of academic publishing can create in the near future for academia. To this aim, I employ the RI lenses for questioning politics and values that lie at the social domain where academic publishing is embedded.

The structure of the article comprises seven sections. After the introduction, a literature review describes the development of the platform economy paradigm. The third section explains the "platformization" of academic publishing and its challenges. The fourth section exposes the role of RI as a political economy approach. The fifth section develops the case studies employed in the essay. The sixth section discusses the findings and the seventh section provides a conclusion for the text.

[2] See https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en#documents.

PLATFORM ECONOMY AS THE IMPERATIVE BUSINESS LOGIC THROUGHOUT DIGITALIZATION

In recent years we have observed how new business models have been spurred into the economy thanks to the quick spread and diffusion of the Internet, the Web and subsequent digital technologies established around these networks such as social media, wearables, cloud computing or artificial intelligence (AI), among others. These technologies have provided to Internet companies with new tools for capturing, collecting, storing, analyzing, treating, reusing and selling data obtained throughout different platforms oriented to facilitate the development of user generated contents (UGC) (Ritzer & Jurgenson, 2010; van Dijck, 2009, 2013).

Digital platforms such as Facebook, YouTube or Uber have been increasingly positioned into daily routines, becoming shortcuts for providing digital services around socializing, streaming or mobility. These examples illustrate the current growing rate of digitalization that society is experimenting, as well as the promotion of data-driven business models. Digital platforms are a key infrastructure in this socio-economic transformation. They combine more visible aspects related with technology, marketing or organizational issues with other previously not really present in business management such as political and media connotations. Platforms seem to be a new buzzword that is loosely defined, and at the same time, a general trend in business (Gillespie, 2010). Digital platforms present features of an horizontal marketplace, as well as maintaining the classical hierarchy of business as usual, what it makes not easy to comprehend their practicalities as a whole (Sundararajan, 2016).

Digital platforms are usually understood as infrastructures based on data collection and classification where data is produced by users' interactions mediated by platforms. The prominence of digital platforms in the actual economy has also drawn the attention of different scholars that have tried to shed some light into their implications (Gillespie, 2010; Gray & Suri, 2019; Srnicek, 2017; Sundararajan, 2016; van Dijck, Poell & Waal, 2018; Zuboff, 2019). Some of these scholars started speaking about a "platform economy" for referring to...

> ...a term that encompasses a growing number of digitally enabled activities in business, politics, and social interaction (Kenney and Zysman, 2016, p. 62).

Major representatives of this new kind of emerging but widely adopted economy (Srnicek, 2017) such as Facebook, Google or Netflix have majorly benefited from initial positive connotations related to digitalization. Some associated ideas to platforms such as sharing economy (Sundararajan, 2016), collaborative consumption (Botsman & Rogers, 2011) and crowdsourcing (Stefano, 2016) have contributed to develop an extensive, complex and meshed socio-technical infrastructure oriented to enable

UGC (Ritzer & Jurgenson, 2010; van Dijck, 2009). But also to capitalize the free availability of digital commons (Fuster-Morell, 2010; Tabarés, 2018) actively created by Internet and Web users.

Other authors refer to this phenomenon of data availability enabled by user activity in different devices, protocols and platforms as "digital labor" (Scholz, 2012) or "free labor" (Terranova, 2000). More recently, other terms such as "data colonialism" (Couldry & Mejias, 2020), "ghost work" (Gray & Suri, 2019) and "surveillance capitalism" (Zuboff, 2019) have been employed for illustrating a new mode of capitalism focused on data, that extracts surplus from user activity in digital platforms (Tabarés, 2021), and it is oriented to the exploitation of commons by platform owners (Fuster-Morell, 2010).

Socio-technical ecosystems established around platforms have also positioned themselves in society as cultural intermediaries, pursuing sustainable business models based on data whilst promoting themselves as champions of freedom of expression (Gillespie, 2010). The influence of the "Californian ideology" (Barbrook & Cameron, 1996) can be easily traced in the promotion of platforms as neutral and egalitarian ecosystems where platform users are supported and treated in an equal way (at least in their term of reference). However, digital platforms are not neutral nor egalitarian ecosystems. Platforms are mediated by algorithms, which are technologies designed to categorize and discriminate data results (Gray & Suri, 2019; Noble, 2018; O´Neill, 2017) and they are also rigidly controlled by platform owners and their everchanging terms of reference (Couldry & Mejias, 2020; van Dijck *et al.*, 2018). Algorithms are also critical components of platforms and they constitute very important active assets regarding technological development and economic competitiveness (O´Neill, 2017). In addition, platforms are totally dependent on users contributions for the digitalization of human activities and creating value throughout data generation extracted from social life (Couldry & Mejias, 2020; Kenney & Zysman, 2016; van Dijck *et al.*, 2018).

All in all, the popularization of platforms during the Web 2.0 period (Tabarés, 2018), the consolidation of the social media phenomenon (van Dijck, 2013) and the breakthrough diffusion of mobile devices worldwide (Vogelstein, 2013) have contributed to a dramatic change in business paradigm, favoring a transition to digital services promoted by big technological companies and nascent startups. That is why some authors like Martin Zenney and John Zysman argue that we are witnessing a major reorganization of our economy in which digital platforms are accumulating too much power:

> If the industrial revolution was organized around the factory, today´s changes are organized around these digital platforms, loosely defined. Indeed, we are in the midst of a reorganization of our economy in which the platform owners are

> seemingly developing power that may be even more formidable than was that of the factory owners in the early industrial revolution (Kenney & Zysman, 2016, p. 62).

"PLATFORMIZATION" OF ACADEMIC PUBLISHING, OPEN ACCESS AND THE CRISIS OF SCIENCE

During the next section, I will try to map out how academic publishers have been no exception at all for this new business logic pushed forward by digital platforms. In fact, many of them started its digital transformation several years ago with the objective of updating its infrastructure and making its contents "platform ready" (Helmond, 2015). To this aim, I pay special attention to three factors that are enmeshed into the social fabric of academic publishing industry and that sustains the extractive and exploitative character of the sector. These three main drivers are the business concentration that occurs through an established oligopoly in the sector, the use of metrics and indicators provided by digitalization and the recent introduction of new business models associated to OA. The first of these factors, business concentration, was intimately associated to digitalization, that was a major trend for publishing outlets during the mid-90s and early 2000s. This immediately led to a significant aggregation and congregation of journals among top publishers (Pitt, 2018). These top publishers, namely Elsevier (part of RELX), Springer Nature (formerly Springer-Verlaag), Wiley-Blackwell and Taylor & Francis (part of Informa Group) are also commonly known as "The Big Four", which it can clearly give an idea of its size and market share in the sector. These representatives of the sector accounted for almost 50% of all papers published in 2013 and three of them (Elsevier, Willey-Blackwell and Taylor & Francis) accounted for the 50% of papers published in specific domains such as social sciences during that year (Larivière *et al.*, 2015).

No further recent data has been gathered for this paper but taking a simple look at the benefits that these publishers have declared during most recent years, their self-declared high margins of operation (higher than 30%) (Beverungen *et al.*, 2012; Larivière *et al.*, 2015) and their different acquisitions of services and academic presses (Mirowski, 2018), it can be argued that this congregation of economic power is still growing[3]. In addition, it is also worthy to mention that companies such as Clarivate Analytics (part of Thomson Reuters group) have also seen in digitalization an opportunity to improve, reinforce and to launch new digital services that are at the core of research and research evaluation practices (Aspesi & Brand, 2020). This is the case of ISI Web of Knowledge, the biggest online academic database of peer-

[3] See for instance https://www.publishersweekly.com/pw/by-topic/industry-news/publisher-news/article/78036-pearson-is-still-the-world-s-largest-publisher.html.

reviewed articles that aggregates the Journal Citation of Reports (JCR), Science Citation Index (SCI) and many others, or the Endnote reference manager, which is broadly used in the academic community. These and other services have largely benefited from the rising of digital technologies and their associated data analytics.

Secondly, the increasing dominant position of these players in this sector favored by digitalization and "platformization" drivers has been the subject of many debates because of its lack of transparency in business practices, its adequacy for research practices and its high profit margins based on unpaid labor carried out by academics (Buranyi, 2017; Pirie, 2009). In fact, the business of academic publishing can be considered as one of the test beds for platform capitalism (Srnicek, 2017), as it is based on the voluntary contribution of thousands of academics worldwide for generating content and reviewing it without any payment involved. This "free labor" (Terranova, 2000) is conducted by academics by the sole objective of advancing in their careers towards the recognition of the academic community and peers to their work, as well as the acquisition of merits that can be acknowledged by research and educational institutions, funding agencies, academic communities and others related. As recent studies have shown, the pressure for publishing is high and widely dispersed at all stages of the career (van Dalen, 2021). The extractive and exploitative character of the academic publishing industry towards academics (Beverungen *et al.*, 2012) seems to be also backed up by the wide use of key performance indicators (KPIs) in research career assessments and university rankings (Aspesi & Brand, 2020).

It is important to stress that KPIs and other metrics favored by digitalization and "platformization" , such as the h-index or the journal impact factor, are provided by one of these top players (Clarivate) (Fox, 2020).The use of these KPIs by for-profit publishers in their top-rated journals is one of the backbones of the business model. It helps to attract researchers to these journals, contributing and publishing in these journals, as well as positioning these journals as prestigious publishing outlets. However, the prestige associated to these specific KPIs developed by for-profit publishers thanks to digital platforms is never a satisfying issue that is carved in stone for the academic community (Ferretti *et al.*, 2018; Rafols *et al*, 2012). Popular initiatives promoted during the last years have strongly fight back this notion of prestige associated to these metrics developed by commercial publishers and its algorithms. One of the most renowned initiatives is The San Francisco Declaration on Research Assessment (DORA)[4] that aims to ban the use of these metrics into research evaluation and research funding processes.

[4] https://sfdora.org/.

Digitalization was pre-supposed to revert this situation, due to the new possibilities that the Internet could bring to the academic publishing process (reduction of associated costs to printing processes, new business models not based on subscription, etc.), but in fact it has been quite the opposite (Aspesi & Brand, 2020; Larivière *et al.*, 2015). For sure, significant successful initiatives such as the "First Monday"[5] Journal or the efforts carried out under the Public Knowledge Project[6] (PKP) which made possible and accessible a significant number of open-source resources have demonstrated that digitalization can be a transforming force for the academic publishing industry. Nevertheless, these kinds of examples have been the exception, not the rule. The ongoing processes of congregation and aggregation (and the launching of new journals thanks to digital technologies) of journals have been exacerbated during the last years thanks to digitalization in liaison to the increasing importance of data analytics and impact KPIs of publications and journals facilitated by digitalization.

Last, OA was meant to bring a transformative change within the industry and significant efforts have been promoted during the last years such as the 4S Coalition[7]. This significant initiative has been able to involve a great number of research funding agencies and research agencies across Europe as well as other international organizations such as the World Health Organization (WHO) or the EC. The latter one has pushed forward a great leap forward to the adoption of OA across the Horizon 2020 Framework Programme for Research and Innovation (Delaney *et al.*, 2020) However, OA has not achieved its aims yet, it is still far ahead of them and for-profit publishers have also created new extractive and exploitative business models around it. According to the Open APC initiative the top three academic publishers (Elsevier, Springer and Willey-Blackwell) have largely benefited from the Article Processing Charge (APC) fees that are established when an author publishes an article via the golden route (Burchardt, 2014). Together, these three publishers have a market share of 45,52% of OA fees that at the time that this article is being written sums more than 108 million euros[8] of public funds allocated to cover APCs in OA journals.

OA has been touted as a desirable paradigm to be achieved by research and academic publishing for making scientific information freely available to citizens. In this sense, the COVID-19 outbreak has been a formidable test bed for this claim, with several collectives promoting different campaigns in social media against vaccination,

[5] https://firstmonday.org/ojs/index.php/fm/about.

[6] https://pkp.sfu.ca/about/.

[7] https://www.coalition-s.org/.

[8] See https://treemaps.intact-project.org/apcdata/openapc/#publisher/.

promoting conspiracy theories and suggesting that the virus was designed into a Chinese laboratory. It is in this context, where research activities focused on coronavirus have adopted a more collaborative and experimental approach for making scientific knowledge freely available worldwide (publishing COVID-19 related papers in OA, uploading preprints, etc.), with the goal of speeding up innovation against the virus and promoting scientific education and communication towards citizenship (Fox, 2020). This rapid acceleration in the transition towards OA has been also stressed the role of digital platforms as "information brokers" conferring to some of the major platforms such as Twitter or Facebook the role of "content curators" to distinguish between false and true information regarding COVID-19[9]. At the same time, top academic for-profit academic publishers have also increased their popularity and importance due to their intermediary roles, aggravating its lock-in effects in the academic community (Aspesi & Brand, 2020).

However, recent episodes during the COVID-19 crisis such as in the case of Hydroxychloroquine (Boseley & Davey, 2020) that involved one of the most prestigious journals in medicine, "The Lancet", have also contributed to aggravate the "crisis of science" (Saltelli & Funtowicz, 2017).The crisis of replicability seems to be one of the symptoms of the growing "platformization" of academic publishing (Mirowski, 2018), as well as the serials crisis[10] seem to be also a manifestation of the growing power of digital platforms commanded by this oligopoly. Indeed, the promotion of OA is by no means a controversial issue in the academic community that reflects the increasing lack of trust and transparency that surrounds digitalization of modern science (Tennant, 2018).

RI AS A POLITICAL ECONOMY APPROACH

In this article I promote a critical reflection about the adoption of OA in academic publishing and its consequences in science-society interactions. This article is not interested in stressing the differences between the normative approach of RI (RRI) and its academic approach (RI) (Owen & Pansera, 2019). The aim is to critically engage with the "platformization" of academic publishing and to use RI as a formidable theoretical lenses for unmasking values and politics (Papaioannou, 2020; van Oudheusden, 2014) behind digital innovations and business models pushed forward by academic

[9] See https://www.politico.eu/article/facebook-avaaz-covid19-coronavirus-misinformation-fake-news/.

[10] This term refers to the continuous increase of costs in subscription from libraries to scholarly journals. It is common to observe that the budget of libraries has commonly keep the same or decreased while subscriptions have been on a continuous rising.

publishers. I also try to explore some alternatives that can promote innovation governance into this sector.

In the DNA of RI there is a clear intention of integrating social and ethical concerns into science (Stilgoe *et al.*, 2013; von Schomberg, 2013), as well as proposing deliberative public engagements between innovators and citizens for contributing to innovation governance (van Oudheusden, 2014). Here, it can be argued that several public values that civil society should expect at the forefront of innovation such as accountability, responsibility, sustainability or transparency are not currently well managed and promoted by the companies that compose the oligopoly of for-profit academic publishing (Larivière *et al.*, 2015; Pitt, 2018; Tennant, 2018). In this sense, RI can be a powerful tool for shedding some light in the alternative paths that the development of OA.

At the same time, it is also important to stress that OA is also a component of the broader OS paradigm that aims to update and transform research practices thanks to the potentialities and possibilities that digitalization brings in (Burgelman *et al.*, 2019). While OA aims to make published articles freely available to citizens, OS has a greater ambition for making accessibly early data and research findings shareable to improve and to speed up available knowledge when dealing with societal challenges. In this sense, it can be argued that OS is a step forward to RI, thanks to digital platforms. The benefits of OS are intimately related with the possibilities that digitalization can have for sharing data, information and knowledge across researchers, such as in the case of COVID-19 crisis. It can provide several tools, instruments and resources for disseminating information and knowledge across research communities as well as establishing common research infrastructures that can thrive scientific discoveries (Burgelman *et al.*, 2019). However, OS does not address critically on the different challenges that modern science suffers today such as the democracy deficit, the reproducibility crisis and the increasing distrust on science by public opinion (Mirowski, 2018; Saltelli & Funtowicz, 2017; Stilgoe *et al.*, 2013). OS pays attention to instruments, tools and procedures, but as it has been argued previously in the text, the problems that science-society interactions face today are far from being new at all. This is the main reason for adopting RI as a political economy approach to shed some light on the current challenges and elucidate some possible solutions and alternatives.

In addition to this, I also employ the abundant literature about platform economy, platform capitalism and digital labor previously commented (Couldry & Mejias, 2020; Kenney & Zysman, 2016; Scholz, 2012; Srnicek, 2017; Terranova, 2000) for providing a critical reflection of the adoption of OA and its consequences in science-society interactions. I pay attention to the reproduction of inequalities in academia

thanks to the combination of digitalization processes and OA development. Employing RI as a political economy approach allow to explore digital transformations carried out in the academic publishing sector. To this extent, RI is positioned in the text in the tradition of social constructivist approach of science. In the following section I employ this approach with the help of two cases that exemplify the challenges that face the implementation of OA throughout digital platforms.

CURRENT CHALLENGES AND POSSIBLE ALTERNATIVES

To start with this critical analysis, the text takes stock of previous literature exposed regarding OA. This is why this article accepts the validity of possible alternatives exposed by Beverungen, Böhm and Land (2012) in which they open an scenario for transformation of the academic publishing industry with four possible responses: further development of open access repositories, a fair trade model of publishing regulation, a renaissance of the university presses, and a self-organized open publishing. These four possible responses are somehow common in the literature and several authors have sketched similar paths for OA development (Aspesi & Brand, 2020; Laakso *et al.*, 2011; Pirie, 2009; Van Noorden, 2013).

However, I also share the concerns of Beverungen *et al.* (2012) as they only consider the self-organized open publishing model as the one with more potential to really provoke a significant change in the sector. This is the model that has been chosen by the famous PLOS One journal[11] which was a great success despite it has not been able to change academics mindset about OA nor the industry itself (Van Noorden, 2013). Indeed, prestigious behind top-rated journals belonging to for-profit publishers, evaluation metrics role such as the impact factor and the lack of associationism and coordinated actions between academics are some of the causes that deter or contravene the impact of these initiatives.

Of these factors, it seems that journals prestige is probably the main barrier for moving to alternatives. It can be said that this status of popular journals in academia have been built up throughout different processes of exploitation and appropriation of free work (Terranova, 2000), mainly consisting in writing and reviewing scientific articles. Both activities are not economically rewarded as these are considered to be part of academic's skillset, as well as competences that can grant public acknowledgement of academia. Peer review activities can be occasionally paid by some journals, but this is not the rule and it is widely acknowledged that scholars will

[11] https://journals.plos.org/plosone/s/journal-information#loc-why-researchers-choose-plos-one.

not be paid for writing and reviewing scientific articles. In addition, the pressure for publishing is high and it's present at all stages of academia (van Dalen, 2021). This exploitation and appropriation of free work is at the core of moral values of academia, leading to increase precarity in combination with other factors spurred by recent economic crisis in several countries.

A second important factor contributing to the development of this prestige is the use of arbitrary metrics and KPIs to quantitively, graphically and numerically express the impact of these top-rated journals. Indeed, academic publishing has been one of the most innovative sectors when using algorithms to sort, classify, manage, refer and suggest pieces of information for platform users. It is also important to remind that Sergey Brin and Larry Page, founders of Google, also were inspired by this use of metrics, when developing its famous algorithm called "PageRank"[12] which organizes and classifies massive amount of information on the Internet. Google search engine was heavily influenced by the number of citations that scientific articles receive for identifying its relevance. KPIs such as impact factor, h-index or number of citations that receive a paper are also at the core of the academic community. During last year's different initiatives such as DORA have also tried to avoid the use of these metrics into research evaluation and research career evaluation, but reality seems to be tough to contest. These arbitrary KPIs developed by for-profit companies are still commonly used by several agencies and institutions, as well as academic communities.

A third important factor that also demands attention is the completely different situation and particularities of academic communities in the whole system of science. Here it can be observed that academic communities that have benefited from strong associations with grand resources to be mobilized have resisted much better that the others the endeavor of combining digitalization and OA transition. Specially, in the social sciences and humanities domain due to the importance of local and regional contexts which has deterred the development of international associations (Larivière *et al.*, 2015). Associationism in the academia seems to be an important factor regarding academic publishing, as this can create alternative paths to be taken by self-organized scientists towards the development of their own digital platforms such as in the case of Science[13].

These three main factors exposed lie at the heart of academia and at the same time, secure business models developed by for-profit academic publishers during last

[12] https://en.wikipedia.org/wiki/PageRank.

[13] Science is published by the American Association for the Advancement of Science.

years. OA was pre-supposed to be a transformative force for the industry, but it was not. Instead, it has aggravated the economic externalities of digitalization in some cases making not affordable to researchers with low resources to publish in OA top-rated journals due to APCs, even in rich countries (Burchardt, 2014). This combination of problematics has deterred the implementation of OA and new initiatives and alternatives have been launched recently for trying to revert this situation. In the following, I pay attention to two particular cases: Open Research Europe (ORE) and Libraria.

Open Research Europe

Open Research Europe (ORE) is the OA publishing platform for Horizon 2020 and future Horizon Europe research results. It was officially launched on the 7th of April of 2021[14]. It is oriented to researchers that have taking part in a Horizon 2020 project or that will be taking part in future Horizon Europe calls and want to publish their original works throughout this OA platform. In the dedicated website of the platform it is also stated that:

> All research is welcome and will be published irrespective of the perceived level of interest or novelty; confirmatory and negative results, as well as null studies are all suitable[15].

ORE aims to provide a free and alternative OA platform for European researchers that have been funded throughout a Horizon 2020 or a Horizon Europe grant and that want to publish any research funded under this umbrella programs in an alternative platform to the classical journal platforms. The fields that covers ORE are natural sciences, engineering and technology, medical and health sciences, agricultural and veterinary sciences, social sciences, and humanities and the arts. All of these fields are addressed by the funding calls and work programmes of Horizon 2020 and Horizon Europe.

ORE uses an open research publishing model that consists of an immediate pre-publication of the article submitted to the platform (after pre-publication checks). Then, the preprint version is published, and data deposited can be viewed and cited. Following this stage, reviewers are selected and invited to conduct an open peer review (names of reviewers as well as their reviews and responses of the authors are public and citable). Last, articles that passed peer review are submitted to indexing

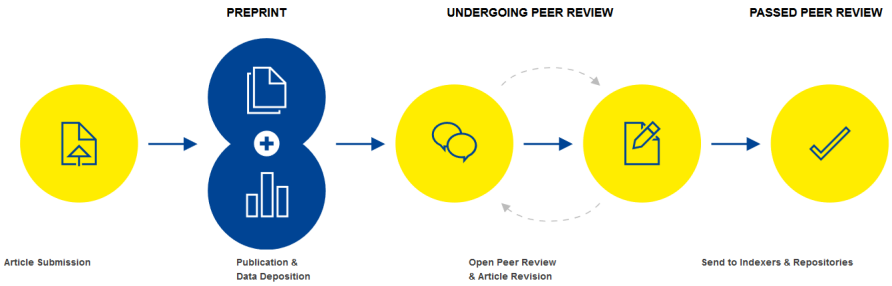
[14] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/other/comm/open-research-europe_official-launch_en.pdf.

[15] https://open-research-europe.ec.europa.eu/about/.

databases and repositories (see figure 1). In a recent interview with Michael Markie, who is the publishing director for F1000[16], the company that provides the technology behind ORE, he stressed how ORE can provide OA to research:

> ORE is an open access publication venue centred around open research practices: open data, open peer review and full transparency of the publication process. It fits into a publishing landscape where there is now real momentum towards full open access to research. Over the last few years, funding bodies have looked to push the envelope with regards to supporting innovation in scholarly communications to ensure the research they fund is open for all to access. As with other funder publishing platforms, ORE provides eligible researchers with an optional venue where they can publish their work and fulfil their open access obligations at no cost to them. (Markie, 2021)

Figure 1 – Publishing process of ORE



In its website, significant benefits are mentioned for researchers, research and society such as no author fees, data sharing, a transparent peer-review, maximizing the value of and impact of EU Research Framework Programmes or:

> ...shifting the way research and researchers are evaluated based on the intrinsic value of the research rather than the venue of publication. (ORE website)

However, at the time that this article is being written there are a number of questions that cannot be answered in the dedicated Q&A section of the website. First of all, it is mentioned that the EC will take over the associated costs for publishing, but it does not provide any information on how much it will be or how it will be funded. It can be assumed that taxpayer's money will be used for these tasks, but no estimation of costs nor annual budgets has been disclosed. Second, it is unclear yet which, how and under which conditions reviewers will be involved for peer reviewing. Reviewers usually agree to conduct reviews in established journals thanks to non-economic incentives such as prestige or access to specific knowledge, but it is still unclear what kind of incentives can offer a platform like this. Last, the launch of ORE has also raised some

[16] https://f1000.com/.

concerns between publishers as the company that provides the technology of ORE, F1000, was acquired by Taylor & Francis in January 2020[17].

Libraria

Libraria is an OA access initiative formed in 2015 by an international group of researchers in the social sciences domain, more specifically on anthropology[18]. This is a recent initiative that tries to promote the subscribe-to-open (S2O) model (Crow, Gallagher & Naim, 2020) into academia for transforming subscription journals to OA. This approach is intended to convert gated access journals to OA using existing library budgets and established relationships. Institutions such as universities, research agencies or technological institutes subscribe to these gated journals in the normal way, and with the assumption that enough benefits are gathered, the journal is published in OA. It is important to stress that this is a subscription model, not a voluntary donation by publishers, that reinforces the relationships between publishers, funding agencies, libraries, researchers and society. In words of some of their promoters:

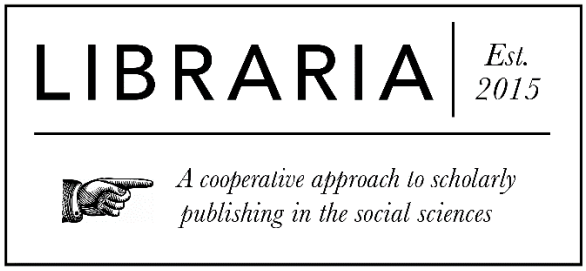
> The major "pro" of this funding model is that it offers a way around a problem currently common to open access publishing — namely, the exploitation of underpaid or volunteer labor of production staff, or of the goodwill of authors and their backing institutions in paying APCs. By escaping proprietary agreements, the L+F model also promises greater budgetary transparency and access to data analytics for all involved. (O'Neill, 2019)

The S2O model enjoys of several benefits that other models don't have such as the APC. For instance, authors are not charged APCs or other administrative fees what it can help to authors from the global south or belonging to modest institutions to publish OA in top-rated journals. The model also provides incentives not only for current subscribers but also for new institutions with lower budgets to join to these collective agreements. The model can be sustainable only if participation is high, what it can balance power relations between publishers and institutions as well as providing future economic sustainability. Yearly renewal processes are similar and do not demand alternative workflows. The S2O model also is characterized by a transitive character and publishers can always start tinkering with some journals instead of its entire collection (Langham-Putrow & Carter, 2020).

[17] See https://newsroom.taylorandfrancisgroup.com/f1000-research-joins-taylor-francis/.

[18] https://libraria.cc/.

Figure 2 – Libraria logo



The model is not free of challenges and these are mainly related with the enough number of subscriptions needed to establishing an economic sustainable model, the strict deadlines between publishers and libraries that demand this subscription model and the common decreases in libraries budgeting that can hinge directly on their subscriptions (Langham-Putrow & Carter, 2020). Another possible challenge for the diffusion of the model outside of non-profit publishers is the alternative funding lines and business models that for-profit models can develop due to the increasing importance of digital platforms, data analytics and AI (Aspesi & Brand, 2020). At the time that this essay is being written, Libraria has helped to reach agreements with publishers such as Annual Reviews, Berghan Books, Coalition Publi.ca and Brill Publishers. Conversations with Oxford University Press and Society for Cinema and Media Studies are underway.

Some observants can argue that this is a similar option to the "transformative deals" (Anderson, 2021) that Elsevier has been negotiating during last years with different institutions around the world, but it is not. First, authors worldwide can benefit from these agreements and publish without APCs. Second, there are more stakeholders involved in the process of negotiation besides the publisher and the institution at stake. Third, there is much more transparency in the deal for all stakeholders affected and access and use of the data analytics of platforms is shared. This favors access to the elaboration and composition of critical indicators that some of the for-profit publishers make use of it for the promotion of diverse top-rated journals. As it can be observed, inclusivity, openness, diversity, transparency and accessibility are some of the values that are promoted by this model.

RESPONSIBILITY IN ACADEMIC PUBLISHING

As it has been stressed, the implementation of OA has faced many struggles and it cannot be said that their main aims have been achieved. With the help of the cases of ORE and Libraria I have tried to expose which kind of barriers are currently confronted by OA initiatives. These barriers are mainly related with the politics and values that lie at the core of different academic communities behind "science". In this sense, and throughout the lenses of RI, OA is not only a matter of making freely accessible to citizens scientific articles. It is also about contesting these politics and values for introducing societal concerns, expectations and public values that can transform current academic publishing.

In this sense, it is of utmost importance to support this transition to other business models where the role of exploitation and appropriation of academic labor do not end up with its commodification and enclosure by representatives of platform economy. OA in this sense can be a transformative force for providing greater visibility to scientific content but also to making this knowledge available and not encapsulated behind paywalls. From a RI perspective it is also important to work against the lack of transparency, responsibility and sustainability that affects the sector. As it is argued by one of the representatives of the two cases discussed, sustainability is one of the main values that should be confronted with the transition to OA.

> The biggest challenge to OA publishing is ensuring sustainable funding. Who will pay the bills to provide free access to knowledge? Publishers, after all, are also in the business of making money. The project of OA is nothing less than to clear a new commons within an economy of publishing that has come, too often, to put profit before science. (O'Neill, 2019)

The serials crisis and the crisis of reproducibility are relatively new phenomena in science but there are logical effects from the increasing digitalization and the rampant pressure for maximizing benefits of platform capitalism (Srnicek, 2017). Academic publishing is no exception to the current business logic imposed by the growing digitalization and the important role of digital platforms on it, but it has been aggravated by their politics and values behind. As Nick Couldry and Ulises Mejias stress, digitalization and datafication try to normalize forms of unpaid and underpaid work that were unthinkable before, but not surprisingly there was a sector where this was previously legitimized.

> Today, social quantification represents the most extensive attempt to construct a whole economy based on the free ride that capitalism can extract from our lives, so that modes of unpaid and underpaid work that were unimaginable before are

> legitimized, normalized, and in the long run, naturalized. (Couldry & Mejias, 2020, p. 58)

For this reason, the role of RI regarding OA should be more ambitious that the current aim of making scientific articles freely available to citizens. It should entail a reconsideration of the current practices and values that are in place behind the whole social domain where academic publishing is embedded. In this sense, it seems clear that several changes can be introduced in the industry for making this sector more oriented to societal concerns and needs. Regulation, investment in community-driven initiatives and public support to associationism in academia for promoting self-OA publishing are some of the directions that should be pursued in the next years.

At the same time and whilst these directions can provide a transformation within the industry, it is also worthy to mention that the full implementation of OA paradigm alone will not solve the challenges previously alluded regarding science-society interactions. Misinformation, public controversies around science and lack of trust in scientists will not be overcome if citizens are still the mere recipients of the increasing and vast scientific production. In this sense, it is so important to start introducing another set of qualitative indicators in research career assessments and research evaluations that can counterbalance current KPIs pushed forward by digital platforms and quantitative indicators commonly accepted in academia. Backing up the efforts of researchers for diffusing, exposing and engaging with citizens around research outputs can be a really transformative force. Public engagement can counterbalance the excessive bias towards impact metrics in academia and can help to strengthen science-society interactions. If no paths are encouraged into this direction I agree with other authors that the transition to OA and OS will met definitively the new configurations enabled by platform capitalism (Mirowski, 2018; Srnicek, 2017).

CONCLUDING REMARKS

As I have explained, challenges faced by academic publishing sector demands a major reorientation of institutional and international initiatives deployed. While a significant focus on OA has been set during the last years through actions such as the 4S Plan, this has also led to reinforce the well-established oligopoly with alternative funding lines via APCs. Here, "platformization" processes conducted by top academic publishers have also helped to augment locked-in effects in their respective "walled gardens". A common side effect of digitalization in many markets (Tabarés, 2021).

In this regard, it is of utmost importance to act in several domains at the same time for progressively transform the sector towards the adoption of OA and the abolition of paywalls. Here, the abundant literature related with digital platforms and digital labor is really relevant for this case as there are a significant number of similarities (Couldry & Mejias, 2020; Scholz, 2012; van Dijck *et al.*, 2018). First, there is an obvious need for international regulation on this sector for creating a common level play of field that can stop abuses from top dominant players. Having an international regulation can help to favor competition as well as favoring the introduction of emergent or incoming players that can contest the well-established oligopoly. Second, there is also a clear need of public investment and backing up of community supported alternatives of OA publishing. Following the same logic of investment that has been pursued during the last years will only reinforce the current dominant position of main players in the sector. The strength of not for profit publishers in certain communities of academia where international associations are not strong, makes a plea for this argument (Larivière *et al.*, 2015). Third, the prominence of metrics in research evaluation and research career assessment need to be counterbalanced with more qualitative indicators encouraging public engagement. Hybrid formats, events or dynamics that can contribute to actively discussing and debating research outcomes with citizens can be probably the most important transformative force.

Last, I would like to stress the limitations of this essay and encouraging researchers to conduct fieldwork and action-research initiatives in this topic. Due to its importance for science-society interactions, it is also surprisingly common to observe the relative limited literature that can be found on the topic as well as the limited awareness and knowledge that can be found on different research communities.

REFERENCES

Anderson, R. (2021). Six Questions (with Answers!) about UC's and Elsevier's New Transformative Deal. *The Scholarly Kitchen*, March 25. https://scholarlykitchen.sspnet.org/2021/03/25/six-questions-with-answers-about-ucs-and-elseviers-new-transformative-deal/

Aspesi, C., & Brand, A. (2020). In pursuit of open science, open access is not enough. *Science*, 368(6491), 574-577. https://doi.org/10.1126/science.aba3763

Barbrook, R., & Cameron, A. (1996). The Californian ideology. *Science as Culture*, 6(1), 44-72. https://doi.org/10.1080/09505439609526455

Beverungen, A., Böhm, S., & Land, C. (2012). The poverty of journal publishing. *Organization*, 19(6), 929-938. https://doi.org/10.1177/1350508412448858

Boseley, S., & Davey, M. (2020). Covid-19: Lancet retracts paper that halted hydroxychloroquine trials. *The Guardian*, June 4. https://www.theguardian.com/world/2020/jun/04/covid-19-lancet-retracts-paper-that-halted-hydroxychloroquine-trials

Botsman, R., & Rogers, R. (2011). *What´s mine is yours. How collaborative consumption is changing the way we live.* London: Harper Collins.

Buranyi, S. (2017). Is the staggeringly profitable business of scientific publishing bad for science? *The Guardian*, June 27. https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science

Burchardt, J. (2014). Researchers Outside APC-Financed Open Access. *SAGE Open*, 4(4), 215824401455171. https://doi.org/10.1177/2158244014551714

Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data*, 2(December), 1-6. https://doi.org/10.3389/fdata.2019.00043

Couldry, N., & Mejias, U. A. (2020). *The Costs of Connection: How Data Are Colonizing Human Life and Appropriating It for Capitalism*. Stanford: Stanford University Press.

Crow, R., Gallagher, R., & Naim, K. (2020). Subscribe to Open: A practical approach for converting subscription journals to open access. *Learned Publishing*, 33(2), 181-185. https://doi.org/10.1002/leap.1262

De Stefano, V. (2016). The rise of the "Just-in-time workforce": On-demand work, crowdwork, and labor protection in the "gig-economy." *Comparative Labor Law & Policy Journal*, 37(3), 471. http://dx.doi.org/10.2139/ssrn.2682602

Delaney, N., Tornasi, Z., Lagher, R., Monachello, R., & Warin, C. (2020). *Science with and for society in Horizon 2020. Achievements and recommendations for Horizon Europe*. Brussels. https://doi.org/10.2777/32018

European Commission (2012). *Responsible Research and Innovation. Europe's ability to respond to societal challenges*. Directorate-General for Research and Innovation. https://doi.org/10.2777/11739

European Commission (2016). *Open innovation, open science, open to the world. A vision for Europe.* Brussels: Publications Office of the European Union. https://ec.europa.eu/digital-single-market/en/news/open-innovation-open-science-open-world-vision-europe

Ferretti, F., Pereira, Â. G., Vértesy, D., & Hardeman, S. (2018). Research excellence indicators: Time to reimagine the "making of"? *Science and Public Policy*, 45(5), 731-741. https://doi.org/10.1093/SCIPOL/SCY007

Fox, J. (2020). Covid-19 Shows That Scientific Journals Need to Open Up. *Bloomberg Opinion*, June 30. https://www.bloomberg.com/opinion/articles/2020-06-30/covid-19-shows-scientific-journals-like-elsevier-need-to-open-up

Fuster-Morell, M. (2010). *Governance of online creation communities: Provision of infrastructure for the building of digital commons*. European University Institute.

Gillespie, T. (2010). The Politics of Platforms. *New Media & Society*, 12(3), 347-364. https://doi.org/10.1002/9781118321607.ch28

Gray, M. L., & Suri, S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global*. Boston: Houghton Mifflin Harcourt.

Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Social Media + Society*, 1(2), 205630511560308. https://doi.org/10.1177/2056305115603080

Kenney, M., & Zysman, J. (2016). The Rise of the Platform Economy. *Issues in Science and Technology*, 32(3), 61.

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B. C., & Hedlund, T. (2011). The development of open access journal publishing from 1993 to 2009. *PLoS ONE*, 6(6), e20961. https://doi.org/10.1371/journal.pone.0020961

Langham-Putrow, A., & Carter, S. J. (2020). Subscribe to Open. *College & Research Libraries News*, 81(1), 18.

Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLoS ONE*, 10(6), e0127502. https://doi.org/10.1371/journal.pone.0127502

Markie, M. (2021). Introducing Open Research Europe (ORE) – Q and A with Michael Markie. *LSE Impact Blog*, March 25. https://blogs.lse.ac.uk/impactofsocialsciences/2021/03/25/introducing-open-research-europe-ore-qa-with-michael-markie/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+feedburner%2FLSEImpactBlog+%28LSE+Impact+Blog%29

Mirowski, P. (2018). The future(s) of open science. *Social Studies of Science*, 48(2), 171-203. https://doi.org/10.1177/0306312718772086

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

O'Neill, K. (2019). Q&A: Heather Paxson on a new model for open-access publishing in anthropology. *MIT News*, May 10. https://news.mit.edu/2019/mit-heather-paxson-open-access-model-workshop-0510

O´Neill, C. (2017). *Weapons of Math Destruction. How Big Data increases inequality and threatens democracy*. London: Penguin Books.

Owen, R., & Pansera, M. (2019). Responsible Innovation and Responsible Research and Innovation. In D. Simon, S. Kuhlmann, J. Stamm & W. Canzle (Eds.), *Handbook on Science and Public Policy* (pp. 26-48). Edward Elgar publishing.

Papaioannou, T. (2020). Innovation, value-neutrality and the question of politics: unmasking the rhetorical and ideological abuse of evolutionary theory. *Journal of Responsible Innovation*, 7(2), 238-255. https://doi.org/10.1080/23299460.2019.1605484

Pirie, I. (2009). The political economy of academic publishing. *Historical Materialism*, 17(3), 31-60. https://doi.org/10.1163/146544609X12469428108466

Pitt, J. (2018). Publish or Impoverish. Academic Publishing and the Platform Economy. *IEEE Technology and Society Magazine*, 37(2), 5-6. https://doi.org/10.1109/MTS.2018.2826058

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy*, 41(7), 1262-1282. https://doi.org/10.1016/j.respol.2012.03.015

Ritzer, G., & Jurgenson, N. (2010). Production, Consumption, Prosumption: The nature of capitalism in the age of the digital "prosumer." *Journal of Consumer Culture*, 10(1), 13-36. https://doi.org/10.1177/1469540509354673

Saltelli, A., & Funtowicz, S. (2017). What is science's crisis really about? *Futures*, 91(May), 5-11. https://doi.org/10.1016/j.futures.2017.05.010

Scholz, T. (2012). Digital labor: The internet as playground and factory. New York: Routledge.

Srnicek, N. (2017). *Platform Capitalism*. Cambridge, UK: Polity Press.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568-1580. https://doi.org/http://dx.doi.org/10.1016/j.respol.2013.05.008

Sundararajan, A. (2016). *The Sharing Economy*. Cambridge, MA: The MIT Press.

Tabarés, R. (2018). Understanding the role of digital commons in the web; The making of HTML5. *Telematics and Informatics*, 35(5), 1438-1449. https://doi.org/10.1016/j.tele.2018.03.013

Tabarés, R. (2021). HTML5 and the evolution of HTML; tracing the origins of digital platforms. *Technology in Society*, 65(May), 101529. https://doi.org/10.1016/j.techsoc.2021.101529

Tennant, J. (2018). Elsevier are corrupting open science in Europe. *The Guardian*, June 29. https://www.theguardian.com/science/political-science/2018/jun/29/elsevier-are-corrupting-open-science-in-europe

Terranova, T. (2000). Free Labor: Producing Culture for the Digital Economy. *Social Text*, 18(2), 33-58. https://doi.org/10.1215/01642472-18-2_63-33

van Dalen, H. P. (2021). How the publish-or-perish principle divides a science: the case of economists. *Scientometrics*, 126(2), 1675-1694. https://doi.org/10.1007/s11192-020-03786-x

van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, Culture, and Society*, 31(1), 41.

van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media.* New York: Oxford University Press.

van Dijck, J., Poell, T., & Waal, M. de. (2018). *The platform society: Public values in a connective world.* New York: Oxford University Press.

Van Noorden, R. (2013). *The true cost of science publishing*. Nature, 495.

van Oudheusden, M. (2014). Where are the politics in responsible innovation? European governance, technology assessments, and beyond. *Journal of Responsible Innovation*, 1(1), 67-86. https://doi.org/10.1080/23299460.2014.882097

Vogelstein, F. (2013). *Dogfight: How Apple and Google Went to War and Started a Revolution*. Sarah Crichton Books.

von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 51–74). London: John Willey & Sons. https://doi.org/10.1002/9781118551424.ch3

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.