Routledge
Taylor & Francis Group

Journal of Responsible Innovation

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjri20

Critical responsible innovation – the role(s) of the researcher

Bernd Carsten Stahl

To cite this article: Bernd Carsten Stahl (2024) Critical responsible innovation – the role(s) of the researcher, Journal of Responsible Innovation, 11:1, 2300162, DOI: 10.1080/23299460.2023.2300162

To link to this article: https://doi.org/10.1080/23299460.2023.2300162

Published online: 25 Jan 2024.

JOURNAL OF RESPONSIBLE INNOVATION
2024, VOL. 11, NO. 1, 2300162
https://doi.org/10.1080/23299460.2023.2300162

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Critical responsible innovation – the role(s) of the researcher

Bernd Carsten Stahl

School of Computer Science, University of Nottingham, Nottingham, UK

ABSTRACT

Critical theory (CT) is a widely used theoretical approach that can be used to describe, understand and justify interventions into scientific and technical research and innovation. It can thus be considered a potential theoretical underpinning of responsible innovation (RI). This article explores the relationship between CT and RI by looking at how CT can influence the practice of researchers who implement RI, typically as part of scientific and technical projects. It proposes a theoretical framework based on three dimensions of criticality: the legitimacy, motivational, and epistemological dimensions. These three dimensions can be used to represent a space of CT where RI scholars can take different roles which influence their RI practice. The article offers some anecdotal evidence to support the validity of the conceptual framework. It proposes a research agenda to validate this framework.

ARTICLE HISTORY
Received 15 January 2023
Accepted 23 December 2023

KEYWORDS
Critical theory; responsible innovation; responsible research and innovation; research agenda

Introduction

Critical theory (CT) is a staple of the social sciences and has been applied to research covering many aspects of modern life, including various different aspects of scientific research and technology development. There are numerous ways of interpreting and defining critical theory. One prominent approach is to see it as a paradigm (Burrell and Morgan 1979; Kuhn 1996), i.e. a broader worldview that can govern many aspects of a scientific endeavour, ranging from epistemology and ontology to eventual methodological choices and interpretations of the meaning of data and evidence (Chua 1986). Most flavours of CT are interested in social interactions, their dynamics, and how these play out to shape relevant societal phenomena.

CT shares this interest in how scientifically and technically enabled social phenomena are created, shaped, and realised with the field of responsible innovation (RI). RI has been described as the attempt to align social needs and preferences with the research and innovation system (Rome Declaration 2014), to ensure that processes and outcomes of research and innovation are socially acceptable, desirable and sustainable. This calls for a detailed understanding of how research and innovation unfold, which dynamics drive them and how these lead to particular outcomes.

CONTACT Bernd Carsten Stahl bernd.stahl@nottingham.ac.uk School of Computer Science, University of Nottingham, Wollaton Road, Nottingham, NG8 1B, UK

Both, RI and CT, make use of methods, theories and approaches stemming from the social sciences and humanities. Both have an ambition to go beyond describing and understanding the phenomena they investigate and to shape them in ways that are ethically sound. This description suggests that there is a potentially significant affinity between CT and RI, allowing them to mutually enforce one another and be employed for their shared purpose. This is, however, not the only possible interpretation. One can alternatively see them as potentially in conflict with CT fundamentally objecting to the market-based constitution of the current research and innovation ecosystem in which RI is typically realised in practice.

The point of highlighting these two different readings of the relationship of CT and RI is to demonstrate that their relationship is not straightforward and subject to debate. This article aims to contribute to this debate by asking how CT can influence RI by shaping the way RI researchers interpret their roles in research projects. Drawing on the literature describing the roles of RI researchers, who are typically social science and humanities (SSH) researchers who are involved in in scientific and technical research, the paper suggests different dimensions of CT (motivation, legitimacy, epistemology) that can influence RI researcher roles. The paper provides anecdotal evidence to suggest that this framework of CT is helpful in describing how RI researchers shape their roles in a large EU research project. It can thus provide the basis for a future research agenda that is spelled out in the conclusion. This is important for the RI discourse and the understanding of possible theoretical underpinnings that inform the way that RI researchers interpret their role and thus drive their practice. More broadly, the article contributes to an understanding of how RI researchers see their role which is important for research leaders, funders and policymakers when considering how best to position RI and how to shape RI integration to promote its broader societal aims.

In order to develop the argument, the paper proceeds as follows. It starts with an overview of its conceptual basis by introducing CT and RI as well as the three dimensions of criticality that are at the centre of the discussion. It then discusses the role of RI researchers in the literature. Subsequently, it introduces a large-scale EU-funded projects to illustrate the CT dimensions. Having demonstrated the relevance of the CT framework, it then concludes by spelling out limitations of the approach and introducing a research agenda that can provide more insights into the topic.

Critical theory and responsible innovation

At the core of this article is the exploration of the relationship between two complex theoretical concepts: RI and CT. CT has a long history and has been employed and developed across many academic disciplines leading to a broad array of interpretations and implementations. RI has a much shorter history but has been a very active field during the last decade, spawning numerous projects aimed at realising the ambitions of RI. RI also draws on conceptual roots that far predate the RI debate in the narrow sense. As a consequence, I will not be able to do justice to all the nuances of the debate of either of these concepts. This section will start with a short overview of CT and highlight three dimensions of criticality, which I believe to be of crucial importance to its relevance

to RI. Using this background of CT, I will then introduce the discourse on RI and the conceptual relationship of CT and RI.

Critical theory

In the continental European usage of the term critical theory is typically traced back to Kant's critiques (Kant 1788; 1995) which provided input into German idealism and the basis for Marx's (1972) critique of capitalism. The various threads of critical theory including the Frankfurt School (Wiggershaus 1995) draw heavily on Marxism in their description of various social phenomena. This is most obvious for first generation members of the Frankfurt School like Horkheimer (1970), Adorno, Marcuse (2002) and to a lesser degree for later thinkers such as Apel (1990), Habermas (1981), Honneth or Forst (2010). One can observe a number of critical theories of different topics, such as critical legal studies (Anon 1987; Unger 1986), critical neuroscience (Choudhury and Slaby 2011) or critical business studies (Alvesson and Deetz 2000; Alvesson and Willmott 2003).

There are prominent accounts of critical theory of technology (Feenberg 1993) which contribute to discussion of CT concerning specific technologies, such as information technology (Brey 2008; Delanty and Harris 2021) or information systems (Brooke 2009; Kvasny and Richardson 2006). Applications of CT to current technology developments continue to be developed, at present most prominently in the field of AI (Krijger 2022; Waelen 2022). The following characterisation of CT relies heavily on sources from these backgrounds.

CT can be understood as a family of approaches that share a number of characteristics (Howcroft and Trauth 2004; Whittle and Spicer 2008). The interpretation of CT as a research paradigm (Orlikowski and Baroudi 1991) hinted at in the introduction can explain how these different aspects fit together and inform the worldview that underlies research activities. At the core of CT as a paradigm is what I have elsewhere called its critical intention (Stahl 2008). By this I mean that CT does not confine itself to describing phenomena in social reality but takes an interventionist stance to change and improve them (Cecez-Kecmanovic 2011). This can be traced to Marx's dictum from his 11th thesis about Feuerbach that philosophers have always interpreted the world, but the important thing is to change it (Marx and Engels 2015). The reason for this critical intention is the recognition that the world is capable of being improved, to facilitate the 'confrontation of the present misery with the objective possibilities hampered by current social conditions' (Frey, Schaupp, and Wenten 2021, 22).

The critical intention to improve social reality can be expressed in many ways. It is often conveyed as the desire to promote emancipation. Emancipation has been called a 'guiding concept of modernity' (Jochum 2021, 30). It has been defined as 'all conscious attempts of human reason to free us from pseudo-natural constraints' (Hirschheim, Klein, and Lyytinen 1995, 83). More simply, emancipation can be understood as meaning 'that more people can achieve their potential to a greater degree' (Klein and Huynh 2004, 163). One important aspect of emancipation in CT is that it not only opposes external social or physical realities but encourages new ways of conceptualising these realities, as Young (2018, 344) puts it, through 'the enactment of new, less oppressive worlds through critical awareness and problem alleviation'. CT thus has a strong

normative component which sets it apart from other research paradigms (Myers and Klein 2011) which tend to be predominantly descriptive. The normativity which informs the desire to change reality requires the motivation to make practical changes. Finally, the epistemological position opens up opportunities to come to different understandings of the world by deconstructing how social reality came about how it is stabilised.

CT usually takes a non-realist or non-objectivist (Alvesson and Willmott 2003) ontological position which means that social reality is not accepted as being 'there' (Lee 1991) but sees it as historically constituted and reproduced by humans. The role of CT is then to deconstruct existing social constructs, understanding how the alienating status quo came into being and offering new perspectives that overcome domination (Myers and Avison 2002). This anti-realist ontology has implications for epistemology, i.e. the question what we can know and what constitutes truth which, in turn, often leads to the preference of qualitative data over quantitative and the reliance on research methods that build on hermeneutics and phenomenology rather than apparently objective positivist approaches.

The emancipatory critical intention is the reason why CT-inspired research tends to focus on obstacles to emancipation. Typical examples of research topics include the mechanisms whereby social realities are removed from scrutiny, for example through ideology (Saravanamuthu 2002) understood as commonplace ideas that are taken for granted as facts or incontrovertible truth (McAulay, Doherty, and Keval 2002). Other topics of continuing interest to CT include power which includes traditional political power in the sense of the capacity to achieve outcomes (Giddens 1984) but also the exertion of power as control through managerial ideologies (Feenberg 1993) or surveillance (Jackson, Gharavi, and Klobas 2006). Power is never just a one-way street but leads to negotiations and resistance (Doolin 2004). It is often linked to particular type of means-ends rationality that shapes how we perceive the world (Kane et al. 2021) and that underpins the modern capitalist approach to life, leading to the commodification of many aspects of life, including work (Greenhill and Wilson 2006) and that promotes the acceptance of a managerialist approach within organisations (Elbanna and Newman 2022).

Much more could be said about CT, but this brief introduction suffices to highlight the conceptual framework that I propose to better understand which roles RI researchers can take. Before we get there, I will need to briefly introduce the concept of RI used in this paper and suggest possible conceptual links to CT.

Responsible innovation

Readers of the journal 'Responsible Innovation' are likely to have a view of what the concept might mean. For the purposes of this article, a broad understanding of the term RI will be used. The EU's take on RI as expressed in the Rome Declaration (2014) emphasises the alignment of social needs and preferences with the research and innovation system. This is consistent with Schomberg's widely cited definition of responsible research and innovation as

> a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on 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 2011, 9)

This, again, is consistent with the definition of RI as a 'collective commitment of care for the future through responsive stewardship of science and innovation in the present' (Owen et al. 2013, 39). For the purpose of this article, the distinction between RRI and RI (Owen, von Schomberg, and Macnaghten 2021) can be ignored, based on the assumption that there are numerous different flavours of RI that have different areas of emphasis and activity but that are collectively consistent in their aim to ensure that science, research and innovation benefit society and that their risks are appropriately addressed.

The following discussion covers areas where this link between CT and RI may be observed or become relevant. This starts with the intellectual provenance of RI. Questions of the relationship of research, science and innovation with society in general have been discussed by various groups and disciplines prior to the emergence of the concept of RI. Some of the more prominent contributors to the RI debate include science and technology studies (STS), technology ethics and philosophy of technology as well as technology assessment (Grunwald 2011). Contributions can also be found from other fields such as legal studies, sociology or philosophy as well as the disciplines that are subject to RI debates, such as nanotechnology, synthetic biology or geoengineering. RI thus has numerous roots and is fundamentally interdisciplinary in nature (Taebi et al. 2014). CT has links to and can inform all of these reference disciplines. STS, for example, as a social science discipline draws on social science theories which include CT. While not all STS scholars would label their work as CT-related, it has been pointed out that CT figures of thought are increasingly relevant to STS and they can mutually support each other (Feenberg 2017). Similarly, not all philosophers and ethicists of technology would consider themselves critical theorists, but there is much work in the philosophy of technology that either explicitly employs CT (Soraker and Brey 2015) or relies on theoretical positions, such as those espoused by members of the Frankfurt School, which are generally seen as representing aspects of CT (Heng and De Moor 2003; Schlagwein, Cecez-Kecmanovic, and Hanckel 2019). In the context of intellectual provenance it is worth highlighting that there are person-related links, for example in the form of Renee Von Schomberg (2020) who is one of the founding figures of RI in the EU and who completed his PhD in political philosophy in Frankfurt. This is an example demonstrating the different possible routes through which CT may have influenced RI. I will call these aspects the dimensions of criticality.

Dimensions of criticality

This section sets out to identify aspects of CT that can have an influence on how researchers see their role in a research project that may have a bearing on the practice of RI. I introduce three such dimensions: the motivational one (the desire to promote change), the normative one (the acceptance of the legitimacy of the status quo) and the epistemological one (the epistemological spectrum from positivism to radical constructivism). These three dimensions were chosen because they are all central to CT, they all form part of the emancipatory agenda.

These three aspects are linked in many ways and have overlaps. They are also all complex composite constructs in their own right. For the purposes of the present argument, they will all be treated as a linear continuum along which an individual can position themselves. This is clearly an ideal-typical reduction, but it can serve as a way of categorising different positions that a researcher can adopt and that has consequences for their relationship with CT and RI. It is useful to provide a brief description of these three dimensions.

The *motivational dimension* focuses on the desire to promote change, in the case of CT this typically implies the desire to further emancipation. The extreme positions along the continuous line of this dimensions would be a purely descriptive attitude, which calls for no change at all and an exclusively interventionist approach which focuses exclusively on promoting change as suggested below:

purely descriptive ↔ purely interventionist

The purely descriptive position is closely aligned with the traditional view of research as description of existing phenomena but probably more wide-spread in the natural science which sometimes do not allow for any intervention (e.g. in astronomy) or intervention only for experimental purposes. In the social sciences there may be a broader recognition that undertaking research can amount to intervention in the phenomenon under study. However, social structures and incentives in academic environments rarely favour an exclusively interventionist position, as this would interfere with traditional academic requirements, such as publication of findings. There are nevertheless approaches that explicitly combine intervention and research, e.g. participatory action research (Argyis and Schon 1989; Whyte 1991). It is thus plausible to think of this dimension as a continuum where an individual scholar can choose how they wish to balance the emphasis on description versus the emphasis on intervention.

The *legitimacy dimension* refers to the researcher's acceptance of the legitimacy of the status quo, i.e. of the socio-economic structures under which the research is undertaken or of the phenomenon under study. Again, one can see this as a continuum where one extreme is the complete acceptance of the legitimacy, and the other extreme is the complete rejection thereof.

acceptance of status quo ↔ rejection of status quo

In practice it is unlikely that one would find many examples of individuals adopting either of these extremes. The full acceptance would imply that one sees the distribution of resources and life chances in current societies as fully justified which is difficult to imagine given the social reality we face in all human societies. The opposite extreme as exemplified by the complete rejection is similarly implausible as it would deny the legitimacy of all aspects of the social order, including its research system. This is a position that an active member of this research system is unlikely to hold. Questions of legitimacy span the entire breadth of research endeavours, ranging from the legitimacy of research questions and research funding to the legitimacy of organisational and power structures within which the research is undertaken to the legitimacy of conclusions drawn from data. They also cover questions of the legitimacy of the RI scholars' position within the research and in the academic system.

The final dimension, the one I have called the *epistemological* one, refers to the researcher's understanding of the nature of research and truth. It is closely interlinked with other aspects of the research paradigm, notably ontology, but I will focus on epistemology for the sake of simplicity. It could be exemplified like this:

strong positivism ↔ anti-positivism

The strong positivist position holds that there can be true statements about an independently existing reality that can be elicited by detached and objective observation, typically represented in quantitative data. The other extreme of anti-positivism, which arguably corresponds to radical constructivism holds that all truth is a social convention based on interaction about phenomena that come about through social interaction. Research has the role of collecting and rendering plausible the collection and synthesis of narratives.

For this dimension it is more difficult to argue that it represents a continuum, as there are categorical differences, e.g. one either believes that truth is an objective description of an independent reality or one does not (Stahl 2007). However, in the social practice of science and research one can observe aspects of a continuum. Critical realism (Callahan 2010; Cruickshank 2010), for example, represents a position that is wedded on the realist underpinnings of positivism but does not fully buy into the epistemological rigour of positivism. The assumption that there may be a continuum on the epistemological dimension is furthermore supported by numerous attempts to bring together aspects of different epistemological positions, often expressed in multi-method approaches that aim to bridge differences between paradigms (Mingers 2001). This is a topic that the paper cannot do justice to, but it may suffice to say that one can conceive of an epistemological continuum at least in the sense that different researchers can attribute plausibility to different positions that may be located along the continuum.

The use of the terms 'positivism' and 'anti-positivism' to denote epistemological positions may raise further questions. Positivism as a position in the philosophy of science can draw on a centuries-long tradition that includes careful discussions that includes questions of limitations of knowledge. I use the term 'positivism' here in the sense of a paradigm of the social sciences as proposed by Burrell and Morgan (1979) and then

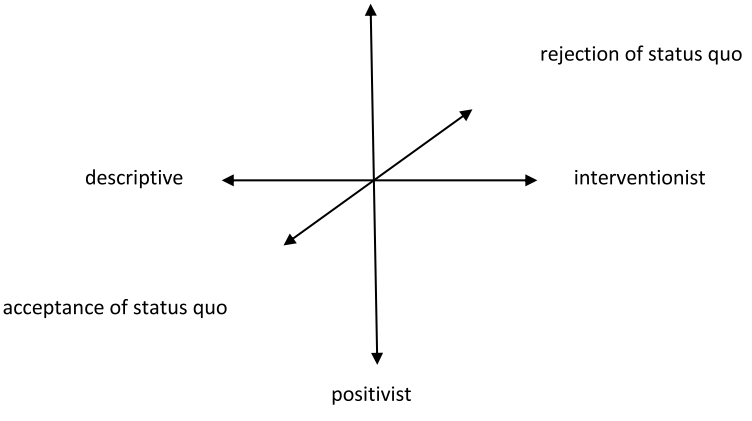

Figure 1. Dimensions of criticality.

developed in various branches of social sciences. This meaning of the word may not do not do justice to sincere positivists, but it has the advantage of being recognisable to many social scientists and having a polar opposite in 'anti-positivism' (Wicks and Freeman 1998), which allows integrating it as an axis in the framework proposed below.

Figure 1 is an attempt to graphically represent the relationship between the different dimensions with a view to expressing the fact that an RI researcher can position themselves anywhere along the continuum of the dimensions of criticality. Representing this space of criticality as a 3-dimensional space should not be misunderstood as implying that the three dimensions are fully independent, nor that these are the only possible dimensions of relevance to CT. It is simply meant to highlight that the adoption of a critical stance is not a simple yes/no decision but requires deliberation of various aspects which, whether intentional or not, will position an RI researcher somewhere in this space.

The way in which Figure 1 reflects the space of criticality, the top, front and right octant would be the space of 'pure' CT, occupied by those who reject the legitimacy of the current capitalist socio-economic model, who see their role mainly as intervening to promote emancipation and who reject the positivist epistemology. The purpose of the figure is to demonstrate that other positions are possible and that they can still lay claim to being based on CT, even though they may diverge from the 'pure' model of CT.

Roles of RI researchers and their link to CT

The idea behind this framework is to find ways to categorise approaches that can inform RI researchers. In principle it is probably possible to undertake RI research from any of the positions represented by the three-dimensional space shown in Figure 1. By locating a researcher in the space represented by Figure 1, it is possible to assess whether and to what degree the three components of CT influence the research approach adopted by RI researchers. This, in turn, gives an indication of the influence that CT has on RI research.

The next step before we come to the way in which researchers relate to the dimensions of criticality is to give an overview of how possible roles of RI researchers have been described in the literature.

Roles of RI researchers in the literature

RI researchers are characterised by the fact that they have a role in planning or implementing RI in the practice of research. They often (but not necessarily) have a background in relevant social science or arts and humanities disciplines, such as science and technology studies or philosophy of technology. This work does not have to be undertaken by someone from outside the core research discipline, but the growing complexity and specialisation of both the scientific and technical research and of RI itself have led to a situation where responsibility for RI is increasingly filled by a specialist.

Possible roles of RI researchers have been discussed for a long time, far predating the current use of the term RI. The traditional public academic intellectual who contributes to broader societal debates represents a well-established example of this type of activity (Foucault 1980). While public intellectuals have long been part of the reflection of science and research, the more wide-spread implementation of RI in projects has led to a growth

in the number and activities of researchers specifically focusing on RI. This has spawned a literature that aims to find ways to describe, categorise and localise the role(s) of RI researchers which is partly aligned with the aims of this article and which is therefore worth reviewing in some more detail.

The role of the RI researcher is sometimes linked to the current research landscape which Balmer et al. (2015) characterise as being post-ELSI, i.e. the next stage after the Ethical, Legal and Social Implications research that started with the Human Genome Project. RI is a key expression of this post-ELSI landscape which is characterised by high levels of interdisciplinarity (Fitzgerald and Callard 2015). Interdisciplinarity is frequently referred to in the discussion of the role of the RI researcher which is not surprising, given that many of them have a background in social sciences or arts and humanities disciplines and they tend to work on projects from the natural science or technology disciplines (Barry, Born, and Weszkalnys 2008).

Interdisciplinary research is generally difficult to implement and raises numerous questions about what is recognised as valid research, how research contributions are recognised and rewarded but also numerous questions about how to do such research, i.e. questions of epistemology and methodology (Freeth and Vilsmaier 2020). There are broadly accepted methods of RI research, such as different flavours of public engagement (Schuijer, Broerse, and Kupper 2021). It is nevertheless worth noting that public engagement is not uncontroversial either and it is not the only way of implementing RI research.

Public engagement is also an example of an activity that can unfold in different types of relationships to the underlying research. It can be used to understand societal needs and thus shape research programmes and calls. It can be employed as part of projects to broaden perspectives, but also to manage risk or solicit acceptance of findings and outcomes. Public engagement can furthermore be used to garner opposition to research activities and divert or stop particular lines of enquiry (Felt and Fochler 2010; Rowe and Frewer 2000; Wynne 2006).

This points to the different relationships that RI researchers can have to the underlying scientific or technical research. Freeth and Vilsmaier (2020) offer an account focusing on the position that RI scholars can adopt. Their experience shows the complexity of this work, and they demonstrate that RI scholars have numerous ways of positioning themselves within the research context. Other attempts to characterise the role of RI scholars focus on specific aspects of the relationship between RI researchers and underlying research, notably on whether RI researchers take a an agonistic–antagonistic stance towards the research (Barry, Born, and Weszkalnys 2008). Another dimension of the role of the RI researcher is whether they focus on observation of the underlying research or whether they practically participate in it (Schuijer, Broerse, and Kupper 2021).

The social reality of taking a position in a research project is more complex than can be described in terms of simple dichotomies (antagonistic – functional or observation – involvement). There have therefore been attempts to paint a richer picture of possible roles. Schuijer, Broerse, and Kupper (2021) propose a taxonomy along two axes (radical – incremental and focus on academic reflection – focus on policy and political action) to identify four different categories of such scholarship (analytic-engaged, analytic-consultative, transformative and pragmatic) and identify five roles that are located across this space (engaged academic, change agent, dialogue capacity builder,

deliberative practitioner, project worker). These roles to some degree correspond with or complement other roles identified in the literature, such as the 'critic' (Balmer et al. 2015; Fitzgerald and Callard 2015) or the 'trickster' (Balmer et al. 2015). The introduction of these dimensions and the roles that RI scholars can play provides a richer picture of social reality and also raises new questions, such as what happens when these different roles come into conflict or different members of a team have different preferred roles. The following Table 1 provides a brief non-comprehensive overview of some of these roles:

This discussion of possible roles of RI scholars provides valuable insights into distribution of activities in research that inspire and guide the following discussion of dimensions of criticality. It is carried by some assumptions that are probably worth unpacking in more detail. One of these seems to be that the RI scholar is an outsider to the community of researchers working on the underlying scientific or technical research. This is certainly true in practice in many funded RI activities but there is no fundamental reason why this has to be the case. The discourse is furthermore somewhat one-sided in its focus on the RI researcher and fails to explore how different roles and responsibilities are conceived and attributed in the underlying research. Interdisciplinary research raises questions not only for RI researchers who have a social science or arts and humanities background but raises very similar questions for natural scientists or technology researchers which may call for the allocation or roles which may be similar in nature to those of the RI researcher with a social science or arts and humanities background.

Having established that there is a recognition of different roles that RI researchers can play in projects, I will now return to the question of how CT can affect or shape those roles and whether there is reason to believe that the dimensions of criticality introduced earlier shape the way RI researchers position themselves. For this purpose, I will first introduce the project from which my observations are drawn and then elaborate on the observations of the relevance of the dimensions of criticality within this project.

Table 1. Roles that (RI) researchers can take, according to the literature.

Role	Description	Reference
Scientific researcher	Learning about a research team with the epistemic goal to create transferable results	(Freeth and Vilsmaier 2020)
Team member	Learning alongside the team	(Freeth and Vilsmaier 2020)
Intervener	Supporting the team to advance its research outcomes	(Freeth and Vilsmaier 2020)
Engaged academic	Observing the current academic system, focused on publications, theory and critique	(Schuijer, Broerse, and Kupper 2021)
Change agent	Advocating for radical transformation, empowering actors to participate	(Schuijer, Broerse, and Kupper 2021)
Dialogue capacity builder	Networking, guiding and training organisations to support the creation of public dialogue	(Schuijer, Broerse, and Kupper 2021)
Deliberative practitioner	Organising and facilitating dialogue aiming to support mutual learning	(Schuijer, Broerse, and Kupper 2021)
Project worker	Setting up and carrying out collaboration with the aim to achieve project goals	(Schuijer, Broerse, and Kupper 2021)
Critic	'Unmasking' scientific development by highlighting interests at play and power relationships	(Balmer et al. 2015)
Trickster	Providing an alternative perspective and disturbing engrained ways of thinking	(Balmer et al. 2015)

RI and dimensions of criticality in the human brain project

In this section I will provide observations of RI practice in the EU-funded Future and Emerging Technologies Human Brain Project (HBP; www.humanbrainproject.eu). This project brought together neuroscience and ICT research. It has 10-year duration (2013–2023) and paves the way for a distributed digital research infrastructure for neuroscience that is called EBRAINS (Amunts et al. 2016). These observations are meant to provide plausibility for the argument that the dimensions of criticality are relevant in RI practice.

One reason why it is an interesting project to illustrate questions concerning RI is that as a very large and interdisciplinary project (more than 500 researchers from more than 100 partner organisations with a core EU funding contribution of more than €400 m) it had a highly visible and large RI presence. I do not want to recount the numerous RI-related activities in the project which have been discussed elsewhere (Rose 2014; Salles et al. 2019; Stahl et al. 2019; Stahl et al. 2021). It is worth noting in the context of this article that the number of RI scholars in the project is significant (>20), that the RI work has been part of the HBP from the outset and that it was highly interdisciplinary including scholars from a range of backgrounds (including philosophy, political science, sociology but also neuroscience, medicine, engineering and archaeology).

While the HBP thus provides an interesting case of RI practice, I should also be clear on the status of the observation offered here. I was part of and held a leadership position in the HBP RI activities from the outset which has provided me with detailed insight into activities and how they came about. The observations as recounted below thus have an aspect of autoethnographic research (Feller and Sammon 2016) but they are not based on a rigorous research protocol, explicit data collection and analysis. They should thus be understood as personal recollections and vignettes that are hopefully interesting and illuminating, but nothing more. It should also be clear that these are my personal views which neither claim to reflect the views of the other RI scholars in the project, nor the official view of the project or of the funder.

The following subsections briefly highlight observations that shed light on the relevance of the dimensions of criticality within the HBP RI work.

Legitimacy dimension

With regards to the legitimacy dimension, it is important to realise that this is in itself a multi-dimensional scale. To give a simple example, a scholar may have different views of the legitimacy of the research system they find themselves in (e.g. publicly funded research in a democratic state), the legitimacy of the project set-up (e.g. a bottom-up consortium assembled to bring together leading experts) or the research topic (e.g. the broadly agreed-upon need to better understand the brain to help individuals with brain-related diseases). The project set-up may result from political considerations leading to the inclusion or exclusion of research expertise based on geographical representation. The research question may be highly contested or lead to easily predictable misuse. Based on these and many other similar considerations, the RI researcher will typically form a view of how legitimate the project as a whole is which will shape the role they wish to play in it.

The legitimacy dimension has influenced the RI work in the HBP in several ways. The legitimacy of the project itself was questioned prominently by members of the neuroscience research community. Early on in the project there was significant disagreement about the management of the project but also about whether it constituted an appropriate investment of significant amounts of public funding which were raised in an open letter published in the journal Nature (Abbott 2015; Frégnac and Laurent 2014). This high-profile debate led to changes of project governance as well as a new scientific focus on developing a distributed digital research infrastructure to support neuroscience. This debate strongly influenced all activities of the project, including the work on RI. A key concern shared by many of the HBP RI groups was not to be seen as providing legitimacy where it was not due – in other words, not to do the 'ethics washing' (Wagner 2018) for the project – but to ensure that RI could contribute to an informed and open debate about the project and its activities to ensure that RI's aim of aligning science and social needs.

Questions of legitimacy also influenced RI scholar's engagement with the scientific work of the project. One topic of debate was the link between neuroscientific research and the resulting understanding of the brain with mental health. The HBP has always had a focus on medical applications but the debate of the degree to which claims about mental health can be based on neuroscientific knowledge remains debated (Rose and Abi-Rached 2013; Rose and Rose 2012). Questions of legitimacy did thus affect the RI work on a general level, but they also appeared with regards to possible consequences of the work. One concern that the HBP RI community spent much time working on was the possibility of misuse of finding and outcomes and of their use in potentially problematic security or military context, which was explored in the work on dual use (Aicardi et al. 2021).

The perception of legitimacy of the project as a whole clearly has an influence on the role that RI scholars see for their work which is closely linked to the motivational dimension.

Motivational dimension

The motivational dimension covers the extent to which an RI scholar wishes to either describe or intervene in the underlying research which will in practice be closely linked to the normative one. However, in principle, it is entirely conceivable for an RI scholar to locate themselves on either end of the scale and see their task as a purely descriptive one or to focus exclusively on changing the course of the research or its impacts. In light of the fact that RI has an explicit normative component, i.e. that it aims to align research and social needs and preferences, a purely descriptive position is unlikely to achieve this goal. One could, however, argue that the changes in research and innovation practices will be motivated and accelerated by a detailed descriptive account of these practices, thus leaving the RI scholar to focus on description. On the other end of the scale, interventions can take many forms. A prominent activity in RI is that of stakeholder engagement which typically requires the RI scholar to undertake extensive practical tasks and manage and steer engagement activities. While such activities can have an interventionist component to them, they may nevertheless remain peripheral for the overall research and innovation activities.

The HBP RI groups and the individuals within them could be located on different points of the motivational continuum from description to intervention. There was a general recognition that the RI activities should have a practical consequence for the

project as a whole and that intervention was therefore an important aim of the project. However, the individual focuses and the way the work was structured differed greatly. Some activities were geared towards the observation and description of the work undertaken within the HBP (Aicardi and Mahfoud 2022) whereas other activities focused predominantly on external impact generation, such as public engagement activities or policy outreach. In practice the exact location of an activity along the continuum of motivation is very difficult to determine, not least because observations of research practices may very well be hugely influential in shaping future practices, but the mechanisms whereby this may happen are difficult to predict and realise. Anecdotal observations suggest that RI scholars in the HBP had vastly different views on this.

Epistemological dimension

The epistemological dimension, finally, is concerned with the RI researchers' views on what constitutes valid knowledge and how it can be produced which has consequences for the possibility and practice of interdisciplinary collaboration. RI researchers can in principle adopt any epistemological position, depending on their research interest and questions. In practice, however, much current RI research seems to be skewed in the direction of interpretive, qualitative and non-positivist work. There are of course notable counter-examples to this, such as the MoRRI work that aimed to identify metrics and indicators for RI (Ravn, Nielsen, and Mejlgaard 2015) or attempts to map RI to existing discourses, such as the business success, corporate social responsibility or sustainable development goals (Yaghmaei 2018; Yaghmaei and van de Poel 2020). The main focus of RI still remains on non-positivist work which raises interesting questions about the relationship between research in natural sciences, engineering, and technology which tends to rely on rarely questioned positivist assumptions and the RI research accompanying it.

The RI scholars working on the HBP RI programme were fairly consistently located on the anti-positivist side of this continuum. Their work did lead to some outputs that included quantitative indicators, for example with regards to aspects of equality, diversity and inclusion or when demonstrating the distribution of particular views or practices across the project (e.g. Stahl and Leach 2023). However, such work was generally couched in terms of a jointly constructed social reality that can be elucidated using quantitative measures rather than the positivist assumption of an external reality that can be objectively described. This anti-positivist stance is consistent with the predominant orientation of the RI discourse as well as most CT-inspired work. It does, however, show a potential disconnect from the typically more positivist research in both neuroscience and ICT that formed the core of the HBP.

These brief considerations of possible relevance of the dimensions of criticality for RI have shown that the three dimensions of criticality have relevance for and can be observed in RI practice. However, there is no clear position that would require an RI scholar to position themselves on particular positions of the continuum. By choosing positions along the three axes of the three-dimensional framework of CT, RI researchers choose their position vis a vis CT, even if they do not do so consciously or intentionally. The question that this observation raises is whether one can observe clusters or patterns in the distribution of positions that RI researchers take that align with the roles described earlier or that describe further roles with a specific reference to CT. The anecdotal evidence provided here

cannot answer this question, but it supports the contention that such roles may exist and be worthy of further investigation, as I now argue in the conclusion.

Conclusion

The purpose of this article was to understand the relationship between CT and RI and in particular explore the role that CT can have in shaping RI practice. It approached this question by looking at the way and extent to which RI scholars can adopt various aspects of CT, here called dimensions of criticality, and which consequences this may have for their RI work. By offering some observations of the way RI was realised in the HBP I have tried to demonstrate that these dimensions of criticality can be observed in practice and have bearings on RI practice.

These considerations are still at an early stage of development and call for further investigations that I will outline below. However, they provide indications of the importance of the topic area and possible implications. One set of implications arises for the discourse on RI and the self-image of the RI researchers who predominantly implement it. It points to the importance of reflecting on the assumptions and underpinnings of RI research and its practitioners itself. Some may consider themselves researchers in the tradition of CT, while others may reject this label. However, all researchers take implicit or explicit positions along the lines of the dimensions of criticality and it is worth making these explicit. There is unlikely to be agreement on which position is most appropriate, but the clarification of existing positions may help communication among RI scholars as well as between RI scholars and the scientists and technologists whose work they accompany. This consideration can thus be translated as a call for a more explicit reflection of the underlying assumptions of RI theory and practice.

A discussion of the role of CT in the field of RI can furthermore contribute to a clear understanding of the role and the purpose of the field itself. RI, understood as a field of study, research practice, or maybe even a discipline, is characterised by the multiplicity and diversity of its reference disciplines, i.e. the disciplinary backgrounds of the members of the field. These are drawn from a large number of areas across the social sciences, humanities, but also natural sciences and technical disciplines. CT has vastly diverging levels of status, relevance, and visibility across these areas. Explicit attention to the topics of CT could delineate more clearly which interpretations of RI exist in the community, how they relate to each other and how these positions affect the relation of the RI field to other fields and disciplines. This internal discourse of the RI community is not just important with regards to the self-reflection of its members but also in the broader societal and political discourse on science that drives the development of parameters of research including funding, esteem and thus the future shape of the broader research system(s) of which RI forms a part.

While such reflections are likely to be of most interest to the RI community, they are of practical importance for those who steer RI at more abstract levels of institutions, funding programmes and research policy. The reason why RI is called for were discussed earlier in the review of the RI discourse. The ostensive reason why funders call for RI integration is that they wish to see societal needs and preference reflected in the research system, which is typically justified in terms of democratic legitimacy of the research system and, in the case of publicly funded research, in terms of

legitimate interests of taxpayers as the ultimate funders of research to see their interest considered. However, in practice research funding and policy are often driven by further considerations which are frequently expressed in financial terms and research is portrayed as an investment that will lead to social benefits. Using this perspective, RI can take the role of creator of legitimacy of research conduct and outcomes. An understanding of the influence of CT allows for a more nuanced review of the intentions behind RI and thus of possible practices, serving to avoid a purely instrumental approach to RI.

These considerations of the relevance of the question discussed in this article highlight the importance of the topic but they also show that the ideas developed here need to be followed up. Before I outline the research programme that could follow from these ideas, I state its limitations.

Limitations

In this article I have tried to portray the link between CT and RI, focusing on the dimensions of criticality that drive the roles that an RI scholar can play in RI. The conceptual model based on the legitimacy, motivational and epistemological dimensions can help to develop a more fine-grained understanding of practical roles of RI scholars and the relevance those have in projects and the broader research context. This approach aims to combine theoretical and practical plausibility with practical feasibility. The cost to pay for this is that it neglects some aspects that call for more scrutiny. For example, the three dimensions of CT proposed above are not the only ones of relevance and discussed in the literature. There are, for example, the focus on the totality of phenomena (Orlikowski and Baroudi 1991) including the historical context (Cecez-Kecmanovic, Klein, and Brooke 2008; Chua 1986) or the importance of embodiment (Adam and Kreps 2006) that have been proposed as components of CT. Similarly, there has been significant attention to reflexivity as a core component of CT (Doolin and McLeod 2005; Kvasny and Richardson 2006; Richardson and Robinson 2007) that directly link CT and RI (Grimpe et al. 2020) which may deserve more attention. I would accept such interventions and interpret them as a call to further enrich the model to the point where it promises to strengthen the model.

A further limitation of the approach is that it uses the CT perspective to interpret RI which some RI scholars are likely to welcome, but others may reject. By using an external theoretical perspective, I gain insights into RI as a phenomenon and subject of study, but I cannot claim to represent the view of all RI scholars. This limitation thus calls for translation work into theoretical positions that RI scholars prefer who do not consider themselves adherents of CT. In order for such translation work to be valuable, however, the conceptual model I presented in this article should first be strengthened and demonstrate its empirical relevance, which is the purpose of the research agenda I present in the next section.

A research agenda

The article so far should have provided the basis of plausibility of these considerations but has only offered anecdotal evidence of its relevance. It gives rise to four sets of research questions:

1. Which factors influence the position that an RI scholar takes with regards to the dimensions of criticality?

Anecdotal evidence indicated earlier and observations from a number of RI projects suggest that RI scholars' positions of legitimacy, motivational and epistemological dimensions are influenced by a number of factors such as disciplinary background, employment status, formal role in a project, or individual characteristics. A better understanding of these influence factors could also help shed light on the following questions.

2. Which roles of RI researchers can be observed in practice based on their adoption of CT?

The answer to this question may relate to the roles discussed earlier, where the adoption of a particular role may be caused by or correlated with aspects of CT. This research question would go beyond the existing literature, however, because it provides a different theoretical grounding that can explain at least aspects of these roles.

3. How does the adoption of relevant aspects of CT influence the practical work of RI scholars?

Put differently, if the answer to the previous question is that there are recognisable roles based on CT, then it will be important to understand whether these lead to specific activities that individuals adopting those roles will normally undertake in RI.

4. What are the implications of these roles and typical activities for the practice and outcomes of RI?

This final question targets the way in which the adoption of CT-based roles makes a practical difference, e.g. in terms of project outcomes and impacts or other measurable project qualities, such as their embedding in larger research programmes or contexts.

The current article cannot provide answers to any of these questions, as they will require broader empirical investigations. Instead, it should be understood as the basis of a larger future research programme that will shed light on the role of CT in RI.

This article raises more questions than it answers. It started from the assumption that CT is likely to be relevant to RI and that a clearer understanding of the way in which CT is adopted by RI scholars can help understand which roles they can play and how this influences the practice and impact of RI work. By offering a conceptual framework constituted by the dimensions of criticality and offering some observations of the relevance of these dimensions, the article develops a framework for a more detailed empirical investigation of the role of CT in RI. This framework and research agenda constitute a contribution to knowledge in their own right because they enrich the theory of RI and open avenues for better understanding of RI's background and practice. If RI is to have a future in the research and innovation governance landscape, then it will need to critically investigate its own assumptions, preconditions and practices to understand which roles it can play in this rapidly developing space. The role that an individual RI researcher adopts in this context is probably not the only important variable, but it is one that would be worth understanding in more detail, given that much of the practices and the eventual consequences of RI are engendered by the individuals who fill it with life in research and innovation projects.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Engineering and Physical Sciences Research Council: [Grant Number EP/T022494/1]; European Commission: [Grant Number 101058587, 101057429]; Horizon 2020 Framework Programme: [Grant Number 945539].

Notes on contributor

Bernd Carsten Stahl is Professor of Critical Research in Technology at the School of Computer Science of the University of Nottingham. His interests cover philosophical issues arising from the intersections of business, technology, and information. This includes ethical questions of current and emerging of ICTs, critical approaches to information systems and issues related to responsible research and innovation.

ORCID

Bernd Carsten Stahl http://orcid.org/0000-0002-4058-4456

References

Abbott, A. 2015. "Human Brain Project Votes for Leadership Change." *Nature*. https://doi.org/10.1038/nature.2015.17060.

Adam, A., and D. Kreps. 2006. "Enabling or Disabling Technologies? A Critical Approach to Web Accessibility." *Information Technology & People* 19 (3): 203–218. https://doi.org/10.1108/09593840610689822.

Aicardi, C., L. Bitsch, N. Bang Bådum, S. Datta, K. Evers, M. Farisco, T. Fothergill, et al. 2021. "Opinion on "Responsible Dual Use" Political, Security, Intelligence and Military Research of Concern in Neuroscience and Neurotechnology." Accessed March 8. https://zenodo.org/record/4588601#.YFW7_q_7SUk.

Aicardi, C., and T. Mahfoud. 2022. "Formal and Informal Infrastructures of Collaboration in the Human Brain Project." *Science, Technology, & Human Values*.

Alvesson, M., and S. A. Deetz. 2000. *Doing Critical Management Research*. London: Sage Publications Ltd.

Alvesson, M., and H. Willmott. 2003. *Studying Management Critically*. London: Sage Publications Ltd.

Amunts, K., C. Ebell, J. Muller, M. Telefont, A. Knoll, and T. Lippert. 2016. "The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain." *Neuron* 92 (3): 574–581. https://doi.org/10.1016/j.neuron.2016.10.046.

Anon. 1987. *Critical Legal Studies*. Oxford: Blackwell.

Apel, K.-O. 1990. *Diskurs Und Verantwortung.: Das Problem Des Übergangs Zur Postkonventionellen Moral*. Frankfurt: Suhrkamp Verlag KG.

Argyis, C., and D. Schon. 1989. "Participatory Action Research and Action Science Compared." *American Behavioral Scientist* 32 (5): 612–623. https://doi.org/10.1177/0002764289032005008.

Balmer, A. S., J. Calvert, C. Marris, S. Molyneux-Hodgson, E. Frow, M. Kearnes, K. Bulpin, P. Schyfter, A. Mackenzie, and P. Martin. 2015. "Taking Roles in Interdisciplinary Collaborations: Reflections on Working in Post-ELSI Spaces in the UK Synthetic Biology Community." *Science & Technology Studies* 28 (3): 3–25. https://doi.org/10.23987/sts.55340.

Barry, A., G. Born, and G. Weszkalnys. 2008. "Logics of Interdisciplinarity." *Economy and Society* 37 (1): 20–49. https://doi.org/10.1080/03085140701760841.

Brey, P. 2008. "The Technological Construction of Social Power." *Social Epistemology* 22 (1): 71–95. https://doi.org/10.1080/02691720701773551.

Brooke, Carole, ed. 2009. *Critical Management Perspectives on Information Systems*. 1st ed. Amsterdam: Butterworth Heinemann.

Burrell, G., and G. Morgan. 1979. *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*. London: Heinemann Educational.
Callahan, G. 2010. "Critical Realism … or Critical Idealism?" *International Journal of Social Economics* 37 (11): 867–879. https://doi.org/10.1108/03068291011082838.
Cecez-Kecmanovic, D. 2011. "Doing Critical Information Systems Research – Arguments for a Critical Research Methodology." *European Journal of Information Systems* 20 (4): 440–455. https://doi.org/10.1057/ejis.2010.67.
Cecez-Kecmanovic, D., H. K. Klein, and C. Brooke. 2008. "Exploring the Critical Agenda in Information Systems Research." *Information Systems Journal* 18 (2): 123–135. https://doi.org/10.1111/j.1365-2575.2008.00295.x.
Choudhury, S., and J. Slaby. 2011. *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*. 1st ed. Holboken, NJ: Wiley-Blackwell.
Chua, W. F. 1986. "Radical Developments in Accounting Thought." *The Accounting Review* 61 (4): 601–632.
Cruickshank, J. 2010. "Knowing Social Reality: A Critique of Bhaskar and Archer's Attempt to Derive a Social Ontology from Lay Knowledge." *Philosophy of the Social Sciences* 40 (4): 579–602. https://doi.org/10.1177/0048393109340664.
Delanty, G., and N. Harris. 2021. "Critical Theory and the Question of Technology: The Frankfurt School Revisited." *Thesis Eleven* 166 (1): 88–108. https://doi.org/10.1177/07255136211002055.
Doolin, B. 2004. "Power and Resistance in the Implementation of a Medical Management Information System." *Information Systems Journal* 14 (4): 343–362. https://doi.org/10.1111/j.1365-2575.2004.00176.x.
Doolin, B., and L. McLeod. 2005. "Towards Critical Interpretivism in IS Research." In *Handbook of Critical Information Systems Research: Theory and Application*, edited by D. Howcroft, and E. Trauth, 244–271. Cheltenham: Edward Elgar Publishing Ltd.
Elbanna, A., and M. Newman. 2022. "The Bright Side and the Dark Side of Top Management Support in Digital Transformation – A Hermeneutical Reading." *Technological Forecasting and Social Change* 175: 121411. https://doi.org/10.1016/j.techfore.2021.121411.
Feenberg, A. 1993. *Critical Theory of Technology*. New ed. New York: Oxford University Press Inc.
Feenberg, A. 2017. "Critical Theory of Technology and STS." *Thesis Eleven* 138 (1): 3–12. https://doi.org/10.1177/0725513616689388.
Feller, J., and D. C. Sammon. 2016. "On the Road to Trusted Data: An Autoethnography of Community Governance and Decision-Making." *Journal of Decision Systems* 25 (sup1): 182–197. https://doi.org/10.1080/12460125.2016.1187413.
Felt, U., and M. Fochler. 2010. "Machineries for Making Publics: Inscribing and De-Scribing Publics in Public Engagement." *Minerva* 48 (3): 219–238. https://doi.org/10.1007/s11024-010-9155-x.
Fitzgerald, D., and F. Callard. 2015. "Social Science and Neuroscience beyond Interdisciplinarity: Experimental Entanglements." *Theory, Culture & Society* 32 (1): 3–32. https://doi.org/10.1177/0263276414537319.
Forst, R. 2010. "The Justification of Human Rights and the Basic Right to Justification: A Reflexive Approach." *Ethics* 120 (4): 711–740. https://doi.org/10.1086/653434.
Foucault, M. 1980. *Power-Knowledge: Selected Interviews & Other Writings – 1972–1977*. Edited by Gordon Colin. New York: Pantheon Books.
Freeth, R., and U. Vilsmaier. 2020. "Researching Collaborative Interdisciplinary Teams." *Science & Technology Studies* 33 (3): 57–72. https://doi.org/10.23987/sts.73060.
Frey, P., S. Schaupp, and K.-A. Wenten. 2021. "Towards Emancipatory Technology Studies." *NanoEthics* 15 (1): 19–27. https://doi.org/10.1007/s11569-021-00388-6.
Frégnac, Y., and G. Laurent. 2014. "Neuroscience: Where Is the Brain in the Human Brain Project?" *Nature* 513 (7516): 27–29. https://doi.org/10.1038/513027a.
Giddens, A. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity.
Greenhill, A., and M. Wilson. 2006. "Haven or Hell? Telework, Flexibility and Family in the e-Society: A Marxist Analysis." *European Journal of Information Systems* 15 (4): 379–388. https://doi.org/10.1057/palgrave.ejis.3000632.

Grimpe, B., B. C. Stahl, C. Ten Holter, P. Inglesant, G. Eden, M. Patel, and M. Jirotka. 2020. "From Collaborative to Institutional Reflexivity: Calibrating Responsibility in the Funding Process." *Science and Public Policy* 47 (5): 720–732. https://doi.org/10.1093/scipol/scaa038.

Grunwald, A. 2011. "Responsible Innovation: Bringing Together Technology Assessment, Applied Ethics, and STS Research." *Enterprise and Work Innovation Studies* 7: 9–31.

Habermas, J. 1981. *Theorie Des Kommunikativen Handelns*. Frankfurt: Suhrkamp.

Heng, M. S. H., and A. De Moor. 2003. "From Habermas's Communicative Theory to Practice on the Internet." *Information Systems Journal* 13 (4): 331–352. https://doi.org/10.1046/j.1365-2575.2003.00144.x.

Hirschheim, R., H. K. Klein, and K. Lyytinen. 1995. *Information Systems Development and Data Modeling: Conceptual and Philosophical Foundations*. Cambridge: Cambridge University Press.

Horkheimer, M. 1970. *Traditionelle Und Kritische Theorie: 4 Aufsätze*. Frankfurt a.M: Athenäum-Fischer-Taschenbuch.

Howcroft, D., and E. M. Trauth. 2004. "The Choice of Critical Information Systems Research." In *Information Systems Research: Relevant Theory and Informed Practice*, edited by B. Kaplan, D. Truex, D. Wastell, A. Wood-Harper, and J. DeGross, 196–211. Boston: Springer.

Jackson, P., H. Gharavi, and J. Klobas. 2006. "Technologies of the Self: Virtual Work and the Inner Panopticon." *Information Technology & People* 19 (3): 219–243. https://doi.org/10.1108/09593840610689831.

Jochum, G. 2021. "Dialectics of Technical Emancipation—Considerations on a Reflexive, Sustainable Technology Development." *NanoEthics* 15 (1): 29–41. https://doi.org/10.1007/s11569-021-00387-7.

Kane, G. C., A. G. Young, A. Majchrzak, and S. Ransbotham. 2021. "Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants." *MIS Quarterly* 45 (1): 371–396. https://doi.org/10.25300/MISQ/2021/1578.

Kant, I. 1788. *Kritik Der Praktischen Vernunft*. Ditzingen: Reclam.

Kant, I. 1995. *Kritik Der Reinen Vernunft. Neuauflage. Studienausgabe*. Frankfurt a.M: Suhrkamp.

Klein, H. K., and M. Q. Huynh. 2004. "The Critical Social Theory of Jürgen Habermas and Its Implications for IS Research." In *Social Theory and Philosophy for Information Systems*, edited by J. Mingers, and L. P. Willcocks, 157–237. Chichester: Wiley.

Krijger, J. 2022. "Enter the Metrics: Critical Theory and Organizational Operationalization of AI Ethics." *AI & SOCIETY* 37 (4): 1427–1437. https://doi.org/10.1007/s00146-021-01256-3.

Kuhn, T. S. 1996. *The Structure of Scientific Revolutions*. New ed. of 3rd rev ed. Chicago: Chicago University Press.

Kvasny, L., and H. Richardson. 2006. "Critical Research in Information Systems: Looking Forward, Looking Back." *Information Technology & People* 19 (3): 196–202. https://doi.org/10.1108/09593840610689813.

Lee, A. S. 1991. "Architecture as a Reference Discipline for MIS." In *Information Systems Research: Contemporary Approaches & Emergent Traditions*, edited by H.-E. Nissen, H. K. Klein, and R. Hirschheim, 573–592. Amsterdam: North Holland.

Marcuse, H. 2002. *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. 1st ed. London and New York: Routledge.

Marx, K. 1972. *Das Kapital. Kritik der politischen Ökonomie. Erster Band, Buch I: Der Produktionsprozeß des Kapitals. 81972. Unveränderter Neudruck der 1. Auflage 1962*. Berlin: Dietz Verlag.

Marx, K., and F. Engels. 2015. *Die deutsche Ideologie: Thesen über Feuerbach: Kritik der neuesten deutschen Philosophie in ihren Repräsentanten Feuerbach, B. Bauer und Stirner und des … in seinen verschiedenen Propheten*. Berlin: Zenodot Verlagsgesellscha.

McAulay, L., N. Doherty, and N. Keval. 2002. "The Stakeholder Dimension in Information Systems Evaluation." *Journal of Information Technology* 17 (4): 241–255. https://doi.org/10.1080/0268396022000017734.

Mingers, J. 2001. "Combining IS Research Methods: Towards a Pluralist Methodology." *Information Systems Research* 12 (3): 240–259. https://doi.org/10.1287/isre.12.3.240.9709.

Myers, Michael D., and David E. Avison, eds. 2002. *Qualitative Research in Information Systems: A Reader*. London: Sage Publications Ltd.

Myers, M. D., and H. K. Klein. 2011. "A Set of Principles for Conducting Critical Research in Information Systems." *MIS Quarterly* 35 (1): 17–36. https://doi.org/10.2307/23043487.

Orlikowski, W. J., and J. J. Baroudi. 1991. "Studying Information Technology in Organizations: Research Approaches and Assumptions." *Information Systems Research* 2 (1): 1–28. https://doi.org/10.1287/isre.2.1.1.

Owen, R., J. Stilgoe, P. M. Macnaghten, E. Fisher, M. Gorman, and D. H. Guston. 2013. "A Framework for Responsible Innovation." In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, edited by R. Owen, J. R. Bessant, and M. Heintz, 27–50. Holboken, NJ: John Wiley & Sons.

Owen, R., R. von Schomberg, and P. Macnaghten. 2021. "An Unfinished Journey? Reflections on a Decade of Responsible Research and Innovation." *Journal of Responsible Innovation* 8 (2): 217–233. https://doi.org/10.1080/23299460.2021.1948789.

Ravn, T., M. W. Nielsen, and N. Mejlgaard. 2015. "Metrics and Indicators of Responsible Research and Innovation." MoRRI Project Deliverable. http://www.technopolis-group.com/wp-content/uploads/2016/12/2171_D3.2.pdf.

Richardson, H., and B. Robinson. 2007. "The Mysterious Case of the Missing Paradigm: A Review of Critical Information Systems Research 1991–2001." *Information Systems Journal* 17 (3): 251–270. https://doi.org/10.1111/j.1365-2575.2007.00230.x.

Rome Declaration. 2014. "Rome Declaration on Responsible Research and Innovation in Europe." https://ec.europa.eu/research/swafs/pdf/rome_declaration_RRI_final_21_November.pdf.

Rose, N. 2014. "The Human Brain Project: Social and Ethical Challenges." *Neuron* 82 (6): 1212–1215. https://doi.org/10.1016/j.neuron.2014.06.001.

Rose, N., and J. M. Abi-Rached. 2013. *Neuro: The New Brain Sciences and the Management of the Mind*. Princeton: Princeton University Press.

Rose, H., and S. Rose. 2012. *Genes, Cells and Brains: The Promethean Promises of the New Biology*. Reprint ed. New York: Verso.

Rowe, G., and L. J. Frewer. 2000. "Public Participation Methods: A Framework for Evaluation." *Science, Technology, & Human Values* 25 (1): 3–29. https://doi.org/10.1177/016224390002500101.

Salles, A., J. G. Bjaalie, K. Evers, M. Farisco, B. T. Fothergill, M. Guerrero, H. Maslen, et al. 2019. "The Human Brain Project: Responsible Brain Research for the Benefit of Society." *Neuron* 101 (3): 380–384. https://doi.org/10.1016/j.neuron.2019.01.005.

Saravanamuthu, K. 2002. "Information Technology and Ideology." *Journal of Information Technology* 17 (2): 79–87. https://doi.org/10.1080/02683960210145977.

Schlagwein, D., D. Cecez-Kecmanovic, and B. Hanckel. 2019. "Ethical Norms and Issues in Crowdsourcing Practices: A Habermasian Analysis." *Information Systems Journal* 29 (4): 811–837. https://doi.org/10.1111/isj.12227.

Schuijer, J. W., J. Broerse, and F. Kupper. 2021. "Juggling Roles, Experiencing Dilemmas: The Challenges of SSH Scholars in Public Engagement." *NanoEthics* 15 (2): 169–189. https://doi.org/10.1007/s11569-021-00394-8.

Soraker, J., and P. Brey. 2015. "Ethics Assessment in Different Fields – Information Technologies." Deliverable. SATORI Project. http://satoriproject.eu/media/2.b.1-Information-technology.pdf.

Stahl, Bernd Carsten. 2007. "Positivism or Non-Positivism – Tertium Non Datur: A Critique of Philosophical Syncretism in IS Research." In *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, edited by R. Sharman, R. Kishore, and R. Ramesh, 115–142. New York: Springer.

Stahl, B. C. 2008. "The Ethical Nature of Critical Research in Information Systems." *Information Systems Journal* 18 (2): 137–163. https://doi.org/10.1111/j.1365-2575.2007.00283.x.

Stahl, Bernd Carsten, S. Akintoye, L. Bitsch, B. Bringedal, D. Eke, M. Farisco, K. Grasenick, et al. 2021. "From Responsible Research and Innovation to Responsibility by Design." *Journal of Responsible Innovation* 8 (2): 175–198. https://doi.org/10.1080/23299460.2021.1955613.

Stahl B.C., S. Akintoye, B. T. Fothergill, M. Guerrero, W. Knight, and I. Ulnicane. 2019. "Beyond Research Ethics: Dialogues in Neuro-ICT Research." *Frontiers in Human Neuroscience* 13, 105. https://doi.org/10.3389/fnhum.2019.00105.

Stahl, Bernd Carsten, and T. Leach. 2023. "Assessing the Ethical and Social Concerns of Artificial Intelligence in Neuroinformatics Research: An Empirical Test of the European Union Assessment List for Trustworthy AI (ALTAI)." *AI and Ethics* 3 (3): 745–767. https://doi.org/10.1007/s43681-022-00201-4.

Taebi, B., A. Correljé, E. Cuppen, M. Dignum, and U. Pesch. 2014. "Responsible Innovation as an Endorsement of Public Values: The Need for Interdisciplinary Research." *Journal of Responsible Innovation* 1 (1): 118–124. https://doi.org/10.1080/23299460.2014.882072.

Unger, R. M. 1986. *The Critical Legal Studies Movement*. Cambridge, MA: Harvard University Press.

Von Schomberg, R., ed. 2011. *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. Luxembourg: Publication Office of the European Union.

Von Schomberg, R. 2020. "In Memory of Karl-Otto Apel: The Challenge of a Universalistic Ethics of Co-Responsibility." *Topologikno* 26: 151–162.

Waelen, R. 2022. "Why AI Ethics Is a Critical Theory." *Philosophy & Technology* 35 (1): 9. https://doi.org/10.1007/s13347-022-00507-5.

Wagner, B. 2018. "Ethics as an Escape from Regulation: From Ethics-Washing to Ethics-Shopping." In *Being Profiled: Cogitas Ergo Sum*, edited by E. Bayamlioglu, I. Baraliuc, L. A. W. Janssens, and M. Hildebrandt, 84–90. Amsterdam: Amsterdam University Press.

Whittle, A., and A. Spicer. 2008. "Is Actor Network Theory Critique?" *Organization Studies* 29 (4): 611–629. https://doi.org/10.1177/0170840607082223.

Whyte, W. F. E. 1991. *Participatory Action Research*. Thousand Oaks: Sage Publications, Inc.

Wicks, A. C., and R. E. Freeman. 1998. "Organization Studies and the New Pragmatism: Positivism, Anti-Positivism, and the Search for Ethics." *Organization Science* 9 (2): 123–140. https://doi.org/10.1287/orsc.9.2.123.

Wiggershaus, R. 1995. *The Frankfurt School: Its History, Theory and Political Significance*. Cambridge, MA: New ed. Polity Press.

Wynne, B. 2006. "Public Engagement as a Means of Restoring Public Trust in Science – Hitting the Notes, but Missing the Music?" *Community Genetics* 9 (3): 211–220.

Yaghmaei, E. 2018. "Responsible Research and Innovation Key Performance Indicators in Industry: A Case Study in the ICT Domain." *Journal of Information, Communication and Ethics in Society* 16 (2): 214–234. https://doi.org/10.1108/JICES-11-2017-0066.

Yaghmaei, Emad, and Ibo van de Poel, eds. 2020. *Assessment of Responsible Innovation: Methods and Practices*. Abingdon, Oxon; New York, NY: Routledge.

Young, A. G. 2018. "Using ICT for Social Good: Cultural Identity Restoration through Emancipatory Pedagogy." *Information Systems Journal* 28 (2): 340–358. https://doi.org/10.1111/isj.12142.