

Institutionalization of responsible research and innovation in universities seen through system archetypes

Åge Olav Mariussen
Nordland Research Institute, Bodø, Norway, and

Seija Virkkala and Antti Mäenpää
University of Vaasa, Vaasa, Finland

Received 12 August 2021
Revised 18 January 2022
9 September 2022
7 March 2023
20 September 2023
15 August 2024
23 October 2024
Accepted 5 November 2024

Abstract

Purpose – The call for the institutionalization of responsible research and innovation (RRI) has resulted in a broad variety of outcomes in universities. This paper aims to look for explanations for these outcomes through a theory-based typology of archetypes of organizational learning (OL) and power in universities.

Design/methodology/approach – This paper is conceptual and uses theories that have not previously been combined, such as historical institutionalism and OL (exploration and exploitation). Universities aim to be organizations governed by professional administrations. At the same time, some universities are also characterized by academic self-organization by communities of peers. This paper suggests a typology of preconditions for OL through historical institutionalism, which explains power relations driving processes of learning.

Findings – The discussion involves the paradox between two types of OL: administrative organizing and academic self-organization. The hypothetical role of RRI in this paradox is explained by a typology of system archetypes that describe why some universities are expected to learn as organizations and others are not.

Research limitations/implications – The theory-based typology will contribute to future empirical studies and evaluations of policy initiatives, such as RRI, with a conceptual understanding of interrelations between institutional preconditions within universities, processes of OL and the wider systemic networks, clusters and societies surrounding universities.

Originality/value – The system archetypes provide a new perspective on attempts at RRI institutionalization by analyzing OL within the broader context of power relations explained by historical institutionalism.

Keywords Universities, Institutional change, Organizational theory, Learning processes, Organizational change, Responsible research and innovation

Paper type Conceptual paper

1. Introduction

Responsible research and innovation (RRI) is both a scientific and a policy concept. RRI emerged from long discussions in science and technology studies on responsibility and public deliberation as well as on particular concepts, such as anticipatory governance;



ethical, legal and social aspects of technology; and technology assessment (Owen & Pansera, 2019; von Schomberg, 2011). According to RRI, questions of research and innovation turn away from focusing on economic growth or scientific excellence toward a normative orientation defined as “responsibility”. Research and innovation should be value driven, oriented toward basic public values for humanity and have socially desirable outcomes (von Schomberg, 2011). In this way, RRI also has to be understood as a concept applied within the context of politics and administration. RRI was adopted as an EU policy in 2011 and has a focus on societal challenges (von Schomberg, 2011; Stilgoe, Owen, & Macnaghten, 2013; Owen, Pansera, Macnaghten, & Randles, 2021a).

As a policy concept, RRI opens up discussions on institutionalization. Institutionalization of RRI is expected to connect universities, as organizations, more closely to society and societal challenges. However, there have been difficulties translating the emerging theory of RRI into practice in universities (Owen, von Schomberg, & Macnaghten, 2021b; Schuijff & Dijkstra, 2020; Christensen et al., 2020). According to Owen et al. (2021b, p. 226), the institutional impacts of RRI are “patchy and uneven”. Schuijff & Dijkstra (2020) noted that even if RRI practices indicate a variety of values and characteristics, they appear to focus on single RRI aspects, such as stakeholder inclusion and engagement; stimulation of reflection among researchers and innovators; management of ethical, legal and social issues of research; or governance of RRI. Often, some elements of RRI are absorbed into universities’ existing activities (Owen & Pansera, 2019). Partial RRI institutionalization was confirmed in a report from the ETHNA project (Szüdi & Lampert, 2023) studying how organizations that fund and perform research adopt and apply RRI principles.

Universities are different. Some can go in new directions, guided by RRI; others cannot. Researchers have found differences in RRI implementation between universities. According to Ryan, Mejlgaard, & Degn (2021) empirical study of 188 universities, RRI-active universities are larger, more research oriented and more multi-disciplinary oriented than RRI-passive universities. These findings led us to examine the reasons for the differences in adopting RRI policies in universities. Accordingly, our research question is:

RQ1. In what ways and why are institutional preconditions for RRI implementation in universities different?

In this paper, we respond to the research question by building on the existing literature on RRI institutionalization and describe RRI institutionalization in universities using the dimensions of anticipation, reflexivity, inclusion and responsiveness (Stilgoe et al., 2013). Anticipation refers to the potential impacts of RRI projects, both positive and negative, and may reveal new groups of stakeholders not traditionally included in the research process. Reflexivity deals with the ability to reflect on the practicalities of a project and its value systems. Inclusion emphasizes the need for broader participation of stakeholders in the research process. Responsiveness means allowing these stakeholders to affect a project’s overall trajectory, goals and values. Stilgoe et al. (2013) emphasized the connection between the dimensions and their integration as a whole.

There is no generally accepted definition of the notion of RRI (Schuijff & Dijkstra, 2020), but there are many influential conceptualizations of RRI (Ribeiro, Smith, & Millar, 2017). In particular, the term “responsible”, in relation to research and innovation, has been conceptualized in various ways. Overlapping extensions with the notion of sustainability has also created confusion (Wittrock, Forsberg, Pols, Macnaghten, & Ludwig, 2021). This may explain why RRI is often used as an umbrella term (Owen et al., 2021a) interpreted and used in flexible ways. This unclear definition may be one reason why RRI has been difficult to translate into practice in universities (Owen et al., 2021a; Schuijff & Dijkstra, 2020;

[Christensen et al., 2020](#)). The conceptual fragility of RRI was caused by attempts to unify different existing policies and communities under the same umbrella. The result was loose implementation of the concept ([Griessler, Braun, Wicher, & Yorulmaz, 2023](#)).

This paper is conceptual and adds to the theoretical understanding of the differences in the institutionalization of RRI in universities. RRI implementation requires organizational learning (OL) in universities, but universities are specific organizations consisting of organizing (management) and self-organizing (academic staff) elements, which might have different forms of OL, explained here as exploration and exploitation as well as their combinations. We build a typology of RRI institutionalization that combines the diverse ways of OL provided with organizational theory and historical institutionalism to explain the power resources and dynamics between different forms of organizing.

We base our argument on organizing and self-organizing on the findings of [Szüdi & Lampert \(2023\)](#) on RRI uptake in universities. They made a distinction between organizational or top-down leadership (management) and bottom-up organizing from the base (scientific staff) and found different outcomes when it comes to support for or rejection of RRI, both at the leadership and base levels. This means that the challenge of RRI may be responded to differently by university management and scientific staff. We explain how different setups between management and academic staff lead to different paths of RRI institutionalization in universities. We build the typology based on theories such as historical institutionalism and OL, seen as the relation between exploration and exploitation and leave aside the broader empirical investigation of RRI institutionalization in universities ([Forsberg & Wittrock, 2022](#); [Ryan et al., 2021](#)).

We refer to RRI institutionalization processes in universities as possible outcomes of the interaction between leadership (which may consist of administrative management) “organizing” the university, on the one hand, and a more or less “self-organizing” academic staff of scientists and teachers, on the other hand. Management and academic staff may align efforts toward RRI institutionalization, but they may also relate differently to RRI. We refer to the tensions and contradictions between these players that may result in various forms of RRI institutionalization as “paradoxes”. A paradox consists of contradictions between different approaches, such as exploration and exploitation, which can be solved in practice in different ways ([Smith & Lewis, 2011](#)).

To understand the varieties of RRI institutionalization in universities, we are guided by [Hansen, Jensen, & Nguyen \(2020\)](#), who found that the methods underpinning RRI are consistent with systems thinking. Peter [Senge \(2006\)](#) defined “systems thinking” as a discipline for seeing wholes and a view for seeing interrelationships rather than things. He illustrated systems thinking using theory-based system archetypes, which present generative systemic structures as more or less stable outcomes of tensions, interrelationships and processes. This paper uses system archetypes to build a typology of RRI institutionalization in universities. In explaining the different system archetypes of RRI institutionalization, we use the theory of institutional change seen through the prism of different solutions in the relationships between the administrative leadership of a university as both an organization and academic self-organization.

The following section clarifies our research methodology. Section 3 explains the different preconditions for RRI implementation by providing a typology of combinations of weak and strong organizing and self-organizing in universities. It also presents the theory of OL through exploration and exploitation. Section 4 clarifies, with the help of historical institutionalism, how these differences in precondition lead to different processes of institutional changes and RRI institutionalization. Section 5 combines the inputs from organization theory (organizing and self-organizing), OL and institutional changes and

provides university archetypes of RRI institutionalization. In Section 6, we answer the research question and make implications for future research topics.

2. Research design

This paper is theoretical and aims to bring new insight into the types of and reasons for differences in RRI institutionalization in universities. [Jaakkola \(2020\)](#) asserted that theoretical papers consist of domain theories and method theories. A domain theory is characterized by a particular set of knowledge on a substantive topic area situated in a field or domain. A method theory is a conceptual system for studying the constituents of a domain theory. Theoretical papers center on domain theories and the role of a method theory is to provide new insights into a domain theory – for instance, a new explanation of concepts and relationships.

We respond to the research question on the differences in institutional preconditions for RRI implementation in universities and build a typology based on a combination of three *domain theories* and notions: historical institutionalism ([Mahoney & Thelen, 2010](#)), organizing and self-organizing in universities ([Maassen & Stensaker, 2019](#)) and OL through exploration and exploitation ([March, 1991](#)). We build on both the theoretical literature and earlier empirical findings on RRI institutionalization ([Table 1](#)).

RRI is a general concept with several *modi operandi* and forms of existence. This creates certain difficulties in understanding the varieties of RRI institutionalization. Accordingly, we focus on the bigger picture. [Wittrock et al. \(2021, p. 49\)](#) described this need as follows: “If RRI is seen as an integrative approach to the relation between science and society, there is a need to emphasize this overall perspective rather than focusing narrowly on the keys” such as gender equality, public engagement, science education, open access and ethics governance.

This led us to think about the systems perspective between science/university and society, which allows for a flexible look at RRI phenomena. A systems perspective helps in understanding complex systems as a whole ([Arnold & Wade, 2015](#)) and in revealing the connections of its different parts, as well as the connections of these parts outside the organization. [Senge \(2006\)](#) used the notion of systems thinking as a discipline for seeing wholes and a frame for seeing interrelationships rather than things as well as for seeing patterns of change rather than static snapshots. Recent research on systems thinking in

Table 1. Theories and concepts used and their roles

Theory	Main concepts	Authors
Domain theory: historical institutionalism	Institutional change, power	Mahoney & Thelen (2010) ; Streeck & Thelen (2005) ; Thelen (2009)
Domain theory: organizational learning	Exploration, exploitation	March (1991) ; Simon (1991) ; Brix (2019)
Domain theory: organizing vs. self-organizing in universities	Organizing and self-organizing	Krücken & Meier (2006) ; Maassen & Stensaker (2019) ; Frost et al. (2016) ; Szüdi & Lampert (2023)
Method theory: systems thinking	System archetype, system approach	Senge (2006)
Method theory: theory of management of paradox	Paradox (persistent coexistence of tensions), paradox management	Smith & Lewis (2011) ; Hargrave and van de Ven (2017)

Source(s): Authors' own work

learning organizations has referred to relations with stakeholders and actors beyond their own organizational borders (Mak & Hong, 2020; Wesselink & Popa, 2024).

A holistic view emphasizes the connections of universities and their different parts, such as disciplines, with the world outside. When applying systems thinking, we focus on two “parts”: academic staff and management, including their interconnections and the connections outside universities. This disposition leads us to distinguish more generic modes of the varying ways of RRI institutionalization in universities (other than the division of disciplines/faculties).

Senge (2006) illustrated systems thinking through system archetypes that present the generative systemic structure of interrelationships and processes. Archetypes are made up of circles of causality that reinforce and balance feedback loops and delays, making their influence gradual (Senge, 2006, pp. 93–94). Archetypes are systemic, self-reinforcing feedback loops that may create undesired or harmful outputs. They may continue to do so because they are not codified, formalized or easy to observe. Instead, their harmful outputs may emerge in a more or less tacit way, which can only be understood through an analysis of systemic relations.

One example of system archetypes is the “tragedy of the commons” (Senge, 2006, p. 398):

Individuals use a commonly available but limited resource solely on the basis of individual need. At first they are rewarded for using it; eventually, they get diminishing returns, which causes them to intensify their efforts. Eventually, the resource is either significantly depleted, eroded, or entirely used up.

The individuals causing the destruction of the natural resources they are exploiting are not necessarily aware of what they are doing. This tragedy can only be revealed through a closer analysis of its long-term systemic impact on ecology. We apply the notion of the system archetype as a *method theory*, using this methodology to identify the dynamics of self-reinforcing processes that create varieties of RRI institutionalization in universities. To reveal these processes, we need to open up factors that require systemic analysis. Universities are loosely coupled (Weick, 1976); their units are interdependent, but at the same time, they can belong to communities beyond university borders, such as quadruple helix networks composed of academia, government, businesses and civil society (Carayannis & Campbell, 2012) or academic peers. These systemic processes, such as tensions between administrative organization and academic self-organization, can also occur at the network or community level, as well as inside universities. Learning at the university level, with the university seen as an organization, is also determined by power relations between stakeholders. A typology of archetypes is presented in Section 5 of this article.

One might think that the negative impacts of archetypes may be discovered and overcome through system analysis. This is not necessarily so. The second *method theory* used in this article is the theory of management of paradox. According to Smith & Lewis (2011, p. 386), paradox is defined as “contradictory yet interrelated elements that exist simultaneously and persist over time” and is characterized by opposition, interdependence and persistence. Accordingly, the theory of management of paradox emphasizes the persistent coexistence of tensions, which may feed into archetypes. Management theories advise students to accept contradictory elements as simultaneously valid and manage them through a combination of differentiation and synergy, rather than trying to solve the tension between them (ambidexterity) (Hargrave & van de Ven, 2017).

In the paradoxical situation of RRI institutionalization, management and/or scientific staff support and/or resist RRI implementation. The tensions and contradictions between them

may result in various forms of RRI institutionalization. These contradictions between different approaches can be solved in practice in different ways, and this article uses a typology of system archetypes to reveal the different ways of RRI institutionalization in universities. The system encompasses networks beyond an organization, such as a university, and system archetypes are positioned differently when it comes to compositions of organizing and self-organizing in universities. A typology of system archetypes explains how a variety of structures shape and reproduce different processes and outcomes in terms of institutional change in universities, seen both as organizations to be managed and as communities of scientific peers. System archetypes are seen as ideal types and real cases can be mixes of different types. The typology reveals new insights and reduces the complexity of the topic (Jaakkola, 2020).

3. Organizational learning in universities

In looking for preconditions for the institutionalization of RRI in universities, it is relevant to ask whether universities are organizations – in other words, if they are subject to the frameworks of OL or something else, such as communities of peers. As we show below, universities may be very different types of organizations, and they relate in different ways to the external systems in which they may be involved. A common factor is the combination of academic work (teaching, science) and administration.

3.1 Are universities organizations?

A classic vision of a weak university is characterized by a combination of weak administration and teaching that specializes in different, fragmented disciplines. There is no communication across disciplinary barriers. The administration organizes routine teaching and there is a weak or nonexistent focus on research.

There are two competing visions of how a university can become a strong, autonomous organization: through strong, professional administrative knowledge or through shared, cross-disciplinary scientific knowledge (the Humboldtian ideal).

Modern university policy has in various ways promoted administrative knowledge in trying to turn universities into “modern” organizations characterized by “strategic actorhood” (Krücken & Meier, 2006) through enhanced institutional autonomy and by strengthening the authority of institutional – often administrative – leaders and top-down governance practices (Maassen & Stensaker, 2019).

According to Krücken & Meier (2006), universities are *in the process of becoming* organizational actors able to act strategically. Organizational actorhood is closely tied to institutional management and leadership at universities. Krücken & Meier (2006) identified four main elements of a new university model: organizational accountability, mainly through the establishment of evaluation procedures; a tendency for universities to define their “own” organizational goals through mission statements, in which the organizational self is created and openly displayed to others; the elaboration and expansion of formal technical structures around these goals; and the transformation of university management into a profession.

Regarding management approaches to OL in universities, Frost, Hattke, & Reihlen (2016) assumed that the relationship between a university and its environment is managed by a professional administration with a focus on policy goals and indicators, markets for potential university services and internal processes of decision-making made by democratically elected committees. Frost *et al.* (2016) saw the scientific peer communities of academic staff as fragmented in different faculties and disciplines (Becher & Trowler, 2001) and/or as mainly focused on internal issues. According to his view, a scientific community governed by peers is “homegrown” and “within defined”.

An additional element of a modern university is the ongoing differentiation of a fine-grained formal organizational structure. It is equipped with organizational subdivisions for international affairs, personnel development, controlling, gender issues, organizational development, technology transfer, etc. These processes of organizational differentiation are linked to the rise of managerial agency. The professional management of the university is established in parallel with the formal statement of university goals (Krücken & Meier, 2006; Maassen & Stensaker, 2019).

This turn in the direction of administrative knowledge at the core of university organization has led to a decoupling of scientific disciplines inside universities from the Humboldtian idea of scientific knowledge as one cross-disciplinary unit. As pointed out by Maassen & Stensaker (2019), this process turns universities into horizontally decoupled organizations (Weick, 1976).

This decoupling opens up the reemergence of scientific self-organization within disciplines, but this time through systems and networks of innovation moving beyond the borders of the university and connecting to society in various forms of triple or quadruple helix networks based on shared knowledge. The result is an entrepreneurial university that draws strength from various powerful external clusters and networks of innovation.

A university's character as a horizontally decoupled organization hampers OL because there is a lack of communication between the functional units. The scientific knowledge developed by individuals, groups, institutes and faculties may link only to a small extent to learning for the organization as a whole. The knowledge gained in one unit will not inform or be exploited in other parts of the organization. Seen from the perspective of a transnational scientific community, universities are organizational hubs in larger networks. However, it has been argued that OL is needed in universities to analyze the challenges of the external environment, develop strategies and enhance the capacity for change and innovation (Dee & Leisyte, 2017).

According to Dee & Leisyte (2017), the creation and movement of knowledge between units may require the development of new structures in a university, and successful organizational change is a learning process involving both managers and academic staff. Organizational change is challenging to implement due to the lack of a common language, and a shared knowledge base might not accumulate enough OL (Dee & Leisyte, 2017). One way to increase communication and discussion and promote OL is to create cross-disciplinary and cross-unit teams (Dee & Leisyte, 2017; Leisyte, Vilkas, Staniskiene, & Zostautiene, 2017).

3.2 *Organizing and self-organizing in universities*

Historically, the alternative to an administratively led university was the Humboldtian ideal of teaching. Human knowledge was seen as one, with the multidisciplinary university colloquium at its peak. This type of university had a cross-disciplinary scientific community at its core and a board consisting of leading researchers from each faculty. Teaching and research were integrated. Such universities were strong organizations created by academic self-organization. This unit of organizing and self-organizing was based on the idea of the unity of knowledge for societal organization.

The attempt to turn universities into modern organizational actors based on the science of management, thereby copying large corporations, has led to the development of cross-disciplinary forms of scientific self-organization linking a university to wider quadruple helix networks and clusters. Entrepreneurial universities combine wide-reaching scientific industry networks and communities with strong internal scientific disciplines (c.f. Fuchs, Bombaerts, & Reymen, 2023)

We may now use the sometimes paradoxical relationship between organizing and self-organizing in universities to outline four archetypes by making distinctions between strong and weak organizing and self-organizing, whereby administrators and professors have different positions (Table 2). These typologies may be seen as preconditions for RRI implementation.

Universities, when seen as organizations, may be strong in two ways:

- (1) by growing a strong administration and a strong set of specialized scientific communities (the university as a modern “corporation”); or
- (2) through a strong cross-disciplinary community in which scientific leaders are in control (the Humboldtian ideal).

Similarly, universities may survive as relatively weak organizations in two ways:

- (1) with a thin administrative layer managing teaching provided by a fragmented staff;
or
- (2) with an administration serving and promoting a strong set of specialized scientific disciplines deeply integrated in wide-ranging clusters outside (entrepreneurial universities).

Entrepreneurial universities may be weak as organizations, but they may be part of powerful clusters.

We can now unpack this static set of archetypes with a view to the processes that create strengths and weaknesses.

3.3 *Organizational learning toward responsible research and innovation – exploration and exploitation*

Wittrock *et al.* (2021) conducted a comparative study on drivers and barriers of RRI institutionalization and concluded that “National policies, regulatory frameworks, laws and monitoring systems appear to be the most effective drivers” (Wittrock *et al.*, 2021, p. 33). The drivers are most powerful when outside pressures are aligned with intraorganizational measures supporting RRI. In addition, an important driver of RRI institutionalization is matching values and perceptions of organizational identity with one or more RRI aspects. One of the most important barriers to implementing RRI is the perception of science according to the criteria of excellence. There is concern that RRI allows outsiders to intervene in the process of creating new knowledge that leads to the defense of academic freedom (Wittrock *et al.*, 2021). This hinders RRI institutionalization, as the inclusion of stakeholders and public engagement are its key dimensions.

We approach these drivers and barriers through OL and power balances. The structure of formalized organizations tends to shape learning processes. First, organizations learn to reduce risks and uncertainties in their environments (Simon, 1991). In other words, they protect their own existence. Accordingly, March (1991) differentiated between two modes of OL often used in organization theory: exploration and exploitation. Exploration refers to the creation of new knowledge and possibilities, whereas exploitation refers to risk avoidance through the use of existing possibilities and knowledge. OL theory is often built around the processes of exploration and exploitation as well as their balance (Brix, 2019). Classic organizational theory saw organizations as mechanisms protecting against the exploitation of existing methods, procedures and forms of knowledge – in other words, the continuity of the organization. Here, exploitation was contrasted with exploration of the environment to achieve an organization’s adaptability to change – in other words, OL. Too strong exploitation and no

Table 2. Institutional preconditions for RRI: university archetypes based on the paradox of organizing and self-organizing

Self-organizing/organizing in universities	Self-organizing (scientific communities) inside and outside universities
<p><i>Integration of administrative and scientific structures</i></p>	<p><i>Integration of administration, scientific communities and societal actor</i></p>
<p>The strength of the university as an organization (Organizing) <i>Management agency</i></p>	<p>Weak The university is focused on individual teaching within standardized disciplines. Research is weak or nonexistent University policy defines indicators of performance in education</p>
<p>Weak The university is focused on individual teaching within standardized disciplines. Research is weak or nonexistent University policy defines indicators of performance in education</p>	<p>Strong An entrepreneurial university integrated into industrial clusters, with a board consisting of industrial partners. A loosely coupled, open organization with strong self-organizing scientific research communities connected to quadruple helix networks University administration serves scientific communities.</p>
<p>Strong The university as a modern hierarchical organization with institutional autonomy and strong administrative leadership. Horizontal decoupling and competition between scientific disciplines open a space for administrative power</p>	<p>The Humboldtian ideal of cross-disciplinary self-organization is at the core of a strong autonomous university organization. Leading scientists are responsible for administrative tasks. Research and teaching are integrated Administrative staff serve cross-disciplinary academic leadership</p>

Source(s): Authors' own work

exploration might lead to an inability to learn, destruction or replacement. It can be concluded that to survive by adapting to change, strong organizations should combine exploration and exploitation.

Processes of learning occur at different levels of organizations and knowledge transfer and conversion processes occur at the individual, group and organizational levels (Crossan, Lane, & White, 1999). Explorative learning can be seen as a forward mechanism from the individual to the organization, whereas exploitative learning can be seen as a mechanism from the organization to the individual (Crossan *et al.*, 1999; Forsberg & Wittrock, 2022).

Universities, as organizations, and their different units and teams often participate in quadruple helix networks around different topics, such as sustainability, digitalization or energy. We suppose that alignment toward RRI principles and OL and learning from the experience of others occur within quadruple helix networks.

RRI is a flexible notion and the interpretation and implementation of RRI varies across different contexts (Forsberg & Wittrock, 2022). When analyzing and translating the notion of RRI, universities are learning. Within a university, management and scientific staff may relate differently to RRI principles and their relevance. They also translate the notion of RRI and its dimensions differently, which can cause contradictory situations and lead to potential resistance toward the implementation of RRI.

We suppose that there are different setups on explorative versus exploitative learning in universities, such as RRI exploration by individuals and teams (staff) in quadruple helix networks and RRI exploitation explored for the sustainable university brand or other positive brands by management. The other potential setup is that RRI is explored in cooperation between staff and management in quadruple helix networks. The university's official policy encourages employers to explore RRI. Management is active in RRI institutionalization, such as by establishing a new organizational subdivision. The exploitation of RRI can be part of university activities, such as teaching.

However, not only does the notion of RRI change in the process of translation and adoption but the local contexts also need to be changed. The precondition for adopting RRI in universities might require a change in meaning or organizational culture, which means that organizational members must change their mental models and question the way they previously solved problems (Wittrock *et al.*, 2021).

Change driven by powerful actors sometimes ends up in conflict because staff expected to submit to power do not understand what is going on or disagree because they see an intervention as illegitimate. Institutional change in the direction of RRI may be promoted through various forms of dialog, aiming to disseminate a story or account of what RRI is all about. When the understanding of what is going on with RRI is disseminated and accepted, norms may sometimes shift and power relations change. Individual learning may open up a process of OL, whereby new objectives and rules are decided upon and implemented.

OL can lead to institutional changes in norms that facilitate RRI at the organizational level through rule changes. On the other hand, institutional change toward RRI implementation can inform and encourage OL toward exploration through anticipation, reflexivity, inclusion and responsiveness.

Historical institutionalism conceptualizes institutions as balances of power that may sometimes become unbalanced. Institutional change is a result of shifting balances between discretionary power (challengers promoting change) and dominant actors with veto power, thereby promoting continuity (Mahoney & Thelen, 2010). However, the power game identified by historical institutionalism may be blind to the optimal combination of exploration and exploitation identified by classical organizational theory.

In other words, historical institutionalism may help identify cases in which combinations of exploration and exploitation are not achieved. In these cases, as described below, other forms of learning may dominate, but the university, as an organization, is unable to learn.

4. Learning, power and institutional change in universities

Here, we explain RRI institutionalization with the theory of institutional change and relate it to the constellation of organizing and self-organizing in universities (see the previous subsection). Institutional change is a process of OL and adaptation.

Most of the time, institutions and institutionalization are used to explain continuity. In rational choice theory, the emphasis is on institutions as structures that produce equilibrium (North, 1990). Sociological institutionalism focuses on the convergence of practices (DiMaggio & Powell, 1983) and institutionalization often refers to the process of embedding norms, values or modes of behavior in a social system, thereby creating continuity.

As the focus of this paper is to explain differences in RRI institutionalization in universities, historical institutionalism describing varying outcomes of power constellations is more relevant than theories explaining convergence. We connect the discussion of power relations with reference to universities seen as organizations with management and, on the other hand, universities as academically self-organizing. There can be a balance of power, or different forms of organizing can confront each other. We refer to the latter as a paradox of organizing and self-organizing in universities.

Szüdi & Lampert (2023) studied how organizations that perform and fund research adopt and apply RRI principles. In the following, they make a distinction between organizational leadership and bottom-up organizing of the staff (base):

The base means the strength of the organizational structures in the support of the uptake or use of RRI [...] that are launched and managed at various (lower) levels of the organization. As opposed to the leadership dimension, these support measures are mostly started from bottom-up, enabled by the values, awareness, skills and knowledge of the research staff and other organizational stakeholders. (Szüdi & Lampert, 2023, p. 17).

It is possible that on one hand, academic staff want to respond to important societal and environmental challenges, such as RRI, and on the other hand, they may also want to be autonomous and have academic freedom (Carrier & Gartzlaff, 2020). Therefore, RRI sometimes seems to be caught in the tension between these different approaches to organizing. The authors tried to elaborate further on where RRI is situated in this paradox by looking at different elements affecting the balance between organizing and self-organizing in one university.

4.1 Institutional change and historical institutionalism

Varieties of institutional change driven by shifting power relations have been explained by historical institutionalists (Streeck & Thelen, 2005; Thelen, 2009; Mahoney & Thelen, 2010). Change may result from a shift in the balance of power between the discretion (power) of the challenging actors and the veto possibilities (resistance) of the actors defending stability. This creates the typology of institutional change presented in Table 3, in which institutional change is driven by power relations, which can be divided into four outcomes: layering, drift, displacement and conversion.

Layering means the introduction of new rules and activities on top of or alongside existing ones without actually replacing the existing rules or institutions. Instead, existing institutions are revised or combined with new ones. Layering happens when defenders of the status quo prevent a complete turnover of the institution (strong veto possibilities) but are not

Table 3. Typology of power-driven institutional change

Veto possibilities/discretion	Low level of discretion	High level of discretion
Weak veto possibilities	Displacement	Conversion
Strong veto possibilities/resistance	Layering	Drift

Source(s): Based on [Mahoney & Thelen \(2010, p. 19\)](#)

against the introduction of amendments. RRI can be seen as the introduction of a new administrative layer or office but not as a deep transformation of ongoing practices and forms of organization.

Displacement means the removal of existing rules and introduction of new ones in a situation in which resistance is broken down. Displacement within a weak organization may open up chaos, whereby things might start to fall apart.

When both organizing and self-organizing are weak, RRI as a societal challenge might remove existing rules and open up science and universities more for society and sustainability – for instance, as a form of establishing a new university. This is the displacement of the old with something new.

Drift refers to the successful defense of existing rules while overlooking or ignoring challenges in the environment. In such cases, the existing institution is successfully defended, but its basis of existence might be eroded. For instance, management introduces RRI, but scientific staff reject it by appealing for academic freedom. However, the basis of this appeal is eroded when society, including research/university funding criteria, moves toward RRI and sustainability. Drift may result in stability in a situation with a balance of power between two strong opposing forces. It refers to a system with a strong focus on the exploitation of existing solutions.

Conversion is a situation in which institutions are not allowed to decay but instead are redirected toward new goals, functions or purposes. Actors intentionally exploit the inherent ambiguities of an institution, and thus, they can direct it toward new goals and functions. New elites may come to power and organize the shift from within the institution in the direction of their preferred interests ([Mahoney & Thelen, 2010](#)), such as RRI.

As far as we know, historical institutionalism has not previously been applied to explain differences in RRI institutionalization. However, some of its features can be seen in the theory of organizational institutionalism applied by [Owen et al. \(2021a\)](#). According to them, crises, either in the external context or within an organization itself, can create legitimacy challenges to which organizations may respond. Legitimacy challenges can cause internal conflict and fragmentation of shared interpretations of organizational behavior, which results in the reframing or even rejection of an incumbent norm and a process by which the legitimacy of institutionalized organizational practice begins to erode. This can lead to the creation and institutionalization of new practices. Whereas [Owen et al. \(2021a\)](#) described the institutionalization of responsible innovation in universities with multiple co-existing and competing logics, this paper explains the different drivers of RRI institutionalization by drawing from insight into institutional changes due to power constellations.

Examining power relations helps to address the issues of organizing (management) and self-organizing (staff). This occurs when management adopts RRI principles as top-down orders but scientific staff resist them, possibly because they restrict academic freedom, or when staff adopt RRI principles and university management neglect them when emphasizing the fulfillment of the national expectations concerning, for instance, the number of

bachelor's, master's, and doctoral degrees. In the case that both adopt RRI principles and reinforce each other, there is no conflict and no paradox. In the case that both management and staff neglect RRI principles, there is no paradox either regarding RRI institutionalization.

In applying this classification to universities, we need to consider the tension between sometimes competing and sometimes complementary sources of knowledge and power at universities, which leads us to examine RRI institutionalization as a learning process.

5. System archetypes of responsible research and innovation institutionalization in universities

We now summarize our findings and look at possible system archetypes. These archetypes have been formulated by combining the different power balances identified by historical institutionalism with the two dimensions of OL: exploration and exploitation (Table 4).

Archetypes are based on the categories of historical institutionalism. From one point of view, they may be seen as self-reinforcing circles of causation that create stable outcomes. They may also be seen as trajectories of institutional change driven by shifting power constellations between dominant and challenging actors, depending on the dynamics of strength between them. This makes it possible to see the difference between OL based on the tension between exploration and exploitation and other forms of learning that do not take place at the university level, such as learning within disciplines or learning in the cluster surrounding the university.

Archetype 1 – Drift: Resistance to change is too strong, and nothing happens, as both management and academic staff resist or are ignorant of RRI. This is the case of risk aversion, whereby the management, scientific staff and organization learn by doing what they have always done (e.g. teaching) a little better. There is no paradox between management and academic staff in RRI adaptation. Drift is a situation in which a university protects its disciplines and accumulates knowledge at the organizational level following traditional patterns, leading to the rejection of new layers of RRI. The focus of academic staff is on teaching within standardized disciplines, such as replacement. RRI principles at the university level may have no or only minor importance; for instance, inclusion can be seen as

Table 4. System archetypes of RRI institutionalization and organizational learning

Power/learning	Exploration	Exploitation
Drift Strong resistance and strong discretion	No	Lock in
Layering Strong resistance and weak discretion	Organizing Management set up a new RRI office	Organizing
Conversion Weak resistance, strong discretion and shared understanding through communication	Self-organizing Communities of academic peers learn and practice RRI	Self-organizing
Displacement Weak resistance and weak discretion	Lagging University becomes irrelevant for RRI (society) and static universities are closed down	No

Source(s): Authors' own work

the dissemination of scientific findings to the public, and there are neither tools nor motivation for engaging with society.

A too-stable university system that is kept in place by strong contradictory forces often relies on a stable environment. As the environment changes, a drifting university with a limited ability to reform may move one of the alternatives. If the university is a strong institution after all, it may transform into a new system with a new balance of power, but one based on cumulative learning from the old system.

Archetype 2 – Layering: This takes place when resistance is strong and forces of change are weak. This means, for instance, that management want to introduce RRI principles, but academic staff resist. The tension or paradox is managed with layering, which involves hybridization and differentiation, whereby RRI becomes an additional organizational layer. Adding a new layer, for instance, as a specific unit specialized for RRI, might lead to RRI rhetoric with a weak impact on the disciplines. Layering implies the design of management principles at the organizational level. *Anticipation* of opportunities outlines research relevant for responding to societal challenges, which enables the university to *include* diverse stakeholders in research projects.

Archetype 3 – Conversion: This implies a paradox between organizing and self-organizing, as the discretion of RRI adoption is strong, and resistance to adopting RRI is overcome by strong winds of change. Conversion can begin either among self-organizing communities of peers or through the initiative of management (organizing). Paradox leads to OL and change. The new paradigm – implementing RRI in practice – goes beyond the old one, and there is a nonlinear shift. The borders of knowledge will be redefined, and different types of knowledge will be integrated (reflexivity and anticipation). For conversion to take place, it is necessary to include scientists, governments, businesses and civil society activists to implement the new synergy.

Archetype 4 – Displacement: This is where both resistance and discretion are weak forces. In this case, there is no RRI institutionalization paradox. In the case of displacement, the existing university is lagging in RRI exploration, making it irrelevant to wider society. One option is closing down the irrelevant university and establishing a new one according to RRI principles.

In addition to the system archetypes, there is a situation of *consensus* whereby both management and academic staff align with the importance of RRI institutionalization in universities. Neither of these is a dominant or challenging actor in the context of RRI institutionalization. There is no conflict or paradox between management and academic staff; instead, they reinforce each other in adopting RRI principles. According to [Szüdi & Lampert \(2023\)](#), this is a case of strong leadership and a strong base. In such cases, there might have been a long tradition of building institutional support structures (e.g. anticipative governance) and reflecting and adjusting research practices according to societal challenges. In this form of OL, RRI is explored and reflected. The opposite case to the consensus of RRI institutionalization is RRI rejection through mutual understanding by management and academic staff, which leads to drift.

OL is possible in three cases: layering, consensus and conversion. Layering is administrative learning inside a university and conversion involves academic staff. These two forms of learning may well be combined in cases in which the administration and the staff are working together, aligning efforts toward RRI institutionalization through mutual understanding and without conflict. We refer to this system archetype as consensus.

In the case of drift, there is no learning at the university level, but there may be other forms of learning going on, such as inside fixed disciplines. A university characterized by drift is unable to adapt its structure to changes in the environment. It cannot prioritize

resources to set up new disciplines. Accordingly, drift may lead to displacement, as other universities or other types of institutions are set up to fill the changing demands. An example of drift replacement might be an experimental ministry that is able to close down or marginalize static universities that are unable to learn and replace them with something new.

6. Conclusions

This paper began with the question of why and how institutional preconditions for RRI implementation differ in universities. Different preconditions for the processes of RRI institutionalization were classified as system archetypes.

6.1 Why are institutional preconditions different?

We drew the typology of archetypes from the insights of historical institutionalism, which explains institutional changes due to the power constellation between dominant and challenging actors. We identified layering, which is seen as the organizational adaptation most likely implemented via the administration and conversion, whereby scientific staff play an important role. In addition, we found possibilities for consensus between management and academic staff. These are archetypes, where OL based on the interplay between exploration and exploitation is regarded as possible in universities. A colloquium of leading professors or professional staff may reorganize or transform a university as a learning organization.

The introduction of power through historical institutionalism led to the identification of two cases in which universities are not expected to be learning organizations: displacement, whereby the existing university is replaced with something else, and drift, which is the case in which nothing changes, following a lock-in on exploitation. A lack of adaptability (drift) may lead to displacement – for instance, if some R&D ministry closes down an inflexible university. In this case, the learning actor may be the ministry involved in experimental learning through trial and error. Similarly, entrepreneurial universities, which are learning organizations, may be guided by their superior clusters and networks of innovation.

The findings of this paper contribute to the emerging literature on RRI institutionalization in four ways.

First, system archetypes combining learning and power provide a holistic approach to looking at RRI interpretation and translation in universities as an integrated whole, rather than focusing on one feature or element, such as one process dimension or one RRI key (cf. [Schuijff & Dijkstra, 2020](#); [Forsberg & Wittrock, 2022](#)). System archetypes are in line with appeals for an overall perspective on RRI ([Owen et al., 2021b](#); [Wittrock et al., 2021](#), p. 49).

Second, the framework brings different preconditions and varying outcomes of the power constellation between dominant and challenging actors in explaining RRI institutionalization. This differs from the neo-institutionalism, isomorphism and congruence used by [Wittrock et al. \(2021\)](#) and [Ryan et al. \(2021\)](#) in studying RRI institutionalization.

Third, we developed further constellations between leadership (organizing, management) and the base (self-organizing, academic staff) as structural preconditions for RRI institutionalization made by the ETHNA report ([Szüdi & Lampert, 2023](#)) and combined them with the notion of paradox, which will be used in examining universities' responses to the challenge of RRI.

Fourth, the findings contribute to the literature on organizations by looking at different processes of exploration and exploitation in OL in the context of varying types of RRI institutionalization in universities (archetypes). We explained how the paradox caused by the power constellation between university management and more self-organizing academic

staff results in a different mix of exploration and exploitation. The profound institutional changes are illustrated as conversions in the case of RRI institutionalization in universities.

6.2 How do these differences in preconditions shape the process of responsible research and innovation institutionalization?

Interestingly, it would seem that the paradox between organizing and self-organizing could enhance OL. This is most likely due to the fact that some level of tension or resistance opens up possibilities for mutual dialogue. However, one option is that both management and scientific staff agree on RRI, but this does not necessarily lead to better OL, especially if it only leads to rhetoric. Self-organizing academics may challenge conservative management or proactive management may shake up conservative academic staff.

One outcome is conversion. The inclusion of legitimate external stakeholders could be central to this outcome. However, this relies on agents of change able to destabilize a rigid organization with built-in tensions.

An external alliance, such as a cluster, may replace a weak university organization and try to create something new from scratch. Proactive management of conservative staff may be done to introduce RRI as a new element.

The theory-based typology illustrated in Table 4 offers contributions to future empirical studies and evaluations of RRI policy initiatives that aim to enhance OL via societal interaction. It also exposes the tension between self-organizing and organizing and explains various RRI contributions for different types of universities. These findings can be used by managers and researchers as training material on RRI.

Instead of studying the potential differences between disciplines in RRI institutionalization, we focused on the different power constellations between organizing (management) and self-organizing (staff). This can be seen as one limitation of this study.

References

- Arnold, R., & Wade, J. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science*, 44, 669–678.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed.). Buckingham: Open University Press/SRHE.
- Brix, J. (2019). Ambidexterity and organizational learning: Revisiting and reconnecting the literatures. *The Learning Organization*, 26(4), 337–351.
- Carayannis, E., & Campbell, D. (2012). Triple helix, quadruple helix and quintuple helix and how do knowledge, innovation, and environment relate to each other? *International Journal of Social Ecology and Sustainable Development*, 1(1), 41–69.
- Carrier, M., & Gartzlaff, M. (2020). Responsible research and innovation: Hopes and fears in the scientific community in Europe. *Journal of Responsible Innovation*, 7(2), 149–169.
- Christensen, M. V., Nieminen, M., Altenhofer, M., Tangcoigne, E., Mejlgard, N., Griessler, E., & Filacek, A. (2020). What's in a name? Perceptions and promotion of responsible research and innovation practices across Europe. *Science and Public Policy*, 47(3), 360–370.
- Crossan, M. M., Lane, H. W., & White, R. E. (1999). An organizational learning framework: From intuition to institution. *The Academy of Management Review*, 24(3), 522–537.
- Dee, J., & Leisyte, L. (2017). Knowledge sharing and organizational change in higher education. *The Learning Organization*, 24(5), 355–365.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.

- Forsberg, E. M., & Wittrock, C. (2022). The potential for learning from good RRI practices and implications for the usefulness of RRI as an umbrella concept. *The Learning Organization*, 687–712, <https://doi.org/10.1108/TLO-09-2021-0104>.
- Frost, J., Hattke, F., & Reihlen, M. (2016). Multi-level governance in universities: Strategy, structure, control. In J. Frost, F. Hattke, & M. Reihlen (Eds), *Multi-Level governance in universities: Strategy, structure, control*, Higher Education, Dynamics 47, Springer International Publishing, Switzerland, pp. 1–15.
- Fuchs, L., Bombaerts, G., & Reymen, I. (2023). Does entrepreneurship belong in the academy? Revisiting the idea of the university. *Journal of Responsible Innovation*, 10(1), 2208424, <https://doi.org/10.1080/23299460.2023.2208424>.
- Griessler, E., Braun, R., Wicher, M., & Yorulmaz, M. (2023). The drama of responsible research and innovation: The ups and downs of a policy concept. In V. Blok (Ed.), *Putting responsible research and innovation into practice* (pp. 11–33), Springer, Cham, Library of Ethics and Applied Philosophy 40, <https://doi.org/10.1007/978-3-031-14710-4>.
- Hansen, J. Ø., Jensen, A., & Nguyen, N. (2020). The responsible learning organization: Can senge (1990) teach organizations how to become responsible innovators? *The Learning Organization*, 27(1), 65–74.
- Hargrave, T., & van de Ven, A. (2017). Integrating dialectical and paradox perspectives on managing contradictions in organizations. *Organization Studies*, 38(3–4), 319–339.
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26.
- Krücken, G., & Meier, F. (2006). Turning the university into an organizational actor. In Drori, G., Meyer, J., & H. Hwang (Eds), *Globalization and organization: World society and organizational change*, Oxford University Press, Oxford, pp. 241–257.
- Leisyte, L., Vilkas, M., Staniskiene, E., & Zostautiene, D. (2017). Balancing countervailing processes at a Lithuanian University. *The Learning Organization*, 24(5), 327–339.
- Maassen, P., & Stensaker, B. (2019). From organised anarchy to de-coupled bureaucracy: The transformation of university organization. *Higher Education Quarterly*, 73(4), 456–468.
- Mahoney, J., & Thelen, K. (2010). A theory of gradual institutional change In Mahoney J., & Thelen K., (Eds), *Explaining institutional change: Ambiguity, agency, and power* New York, NY: Cambridge University Press, pp. 1–37.
- Mak, C., & Hong, J. (2020). Creating learning organization 2.0: A contextualized and multi-stakeholder approach. *The Learning Organization*, 27(3), 235–248.
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87.
- North, D. C. (1990). *Institutions, institutional change and economic performance. The political economy of institutions and decisions*, Cambridge: Cambridge University Press.
- Owen, R., & Pansera, M. (2019). Responsible innovation and responsible research and innovation. In Simon, D. S., Kuhlmann, J., Stamm, & Canzler, W. (Eds), *Handbook on science and public policy*, Cheltenham: Edward Elgar Publishing, pp. 26–48.
- Owen, R., Pansera, M., Macnaghten, P., & Randles, S. (2021a). Organisational institutionalisation of responsible innovation. *Research Policy*, 50(1), 104132, <https://doi.org/10.1016/j.respol.2020.104132>.
- Owen, R., von Schomberg, R., & Macnaghten, P. (2021b). An unfinished journey? Reflections on a decade of responsible research and innovation. *Journal of Responsible Innovation*, 8(2), 217–233.
- 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(1), 81–103.

-
- Ryan, T. K., Mejlgaard, N., & Degn, L. (2021). Organizational patterns of RRI: How organizational properties relate to RRI implementation. *Journal of Responsible Innovation*, 8(2), 320–337.
- Schuijff, M., & Dijkstra, A. M. (2020). Practices of responsible research and innovation: A review. *Science and Engineering Ethics*, 26(2), 533–574.
- Senge, P. M. (2006). *The fifth discipline: the art and practice of the learning organization*, New York, NY: Doubleday/Currency.
- Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization Science*, 2(1), 125–134.
- Smith, W., & Lewis, M. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of Management Review*, 36(2), 381–403.
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568–1580.
- Streeck, W., & Thelen, K. (2005). Introduction: Institutional change in advanced political economies. In W. Streeck, & K. Thelen (Eds), *Beyond continuity: Institutional change in advanced political economies* Oxford: OUP Oxford, pp. 1–39.
- Szűdi, G., & Lampert, D. (2023). Blueprint for institutional change to implement. *ETHNA D4.3*. Retrieved from https://ethnasystem.eu/wp-content/uploads/2023/03/ETHNA_-D4.3_Institutional_Blueprint_v4.0-final_ngs.pdf
- Thelen, K. (2009). Institutional change in advanced political economies. *British Journal of Industrial Relations*, 47(3), 471–498.
- von Schomberg, R. (2011). *Towards responsible research and innovation in the information and communication technologies and security technologies fields*, Brussels: European Commission.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1–19.
- Wesselink, R., & Popa, E. (2024). Responsible learning organizations: A framework to embed responsible innovation within organizations. *The Learning Organization*, 32(7), 1-17, <https://doi.org/10.1108/TLO-11-2021-0136>.
- Witrock, C., Forsberg, E.-M., Pols, A., Macnaghten, P., & Ludwig, D. (2021). Implementing responsible research and innovation: Organisational and national conditions. *Science and Engineering Ethics*. Retrieved from <https://link.springer.com/book/10.1007/978-3-030-54286-3>

Corresponding author

Seija Virkkala can be contacted at: seija.virkkala@uwasa.fi