Governance of Responsible Research and Innovation: An Agent-Based Model Approach

Harold Paredes-Frigolett\textsuperscript{a,1}, Luiz Flávio Autran Monteiro Gomes\textsuperscript{b} and Javier Pereira\textsuperscript{c}

\textsuperscript{a}Faculty of Economics and Business, Universidad Diego Portales, Av. Santa Clara 797, Huechuraba, Santiago, Chile
\textsuperscript{b}IBMEC School of Business and Economics, Av. Presidente Wilson 118, 20030-020 Rio de Janeiro, Brazil
\textsuperscript{c}Faculty of Engineering, Universidad Diego Portales, Ejército 441, Santiago, Chile

Abstract

Responsible research and innovation is a new research field that is still emerging, particularly in connection with recent efforts within the European Commission. Inspired in these recent developments, we put forth the first agent-based model of responsible research and innovation (RRI). The model has been designed as a decision-aiding tool for both policymakers driving innovation policy and innovation managers facing the complex trade-offs posed by the involvement of civil society organizations in innovation agendas. The model allows policymakers and innovation strategists to inform and guide public policies and strategic decisions regarding the governance of responsible research and innovation.

1. Introduction

RRI is a research field that has gained importance in recent years, as evidenced by the efforts of the European Commission to propose a framework for RRI [1-3]. Although closely connected with the field of corporate social responsibility (CSR) [4], RRI is a research field that has not yet been exposed to public scrutiny in the same way CSR has. In this article, we put forth the first agent-based model of responsible research and innovation (RRI). The model is intended to be used by policymakers and corporations to simulate the effect of different RRI governance models as well as the role of so-called civil society organizations (CSOs hereinafter) in RRI governance. The model will serve as a tool: (i) for the evaluation \textit{ex ante} of the effect of public policies in the area of RRI governance and (ii) for strategic decision-making by innovation strategists. The article is organized as follows. In Section 2, we briefly present the previous work on RRI governance our model is based upon. In Section 3, we present the central research question as well as the hypotheses of the

\textsuperscript{1} Corresponding author. Tel.: +56-2-2213 0127
\textit{E-mail address}: harold.paredes@udp.cl
model. In Section 4, we present the theoretical framework developed. In Section 5, we present the agent-based model. In Section 6, we present our conclusions.

2. Background

In this section, we briefly present the relevant literature that serves as the theoretical basis for our model.

2.1. The Dimensions of RRI

RRI research has been comprised of qualitative empirical research [5-10]. In a recent case study, Stilgoe et al. [10] propose a characterization of RRI governance as consisting of the following dimensions:

1. Anticipating the potential positive and negative outcomes of research and innovation agendas;
2. Reflecting on the internal norms and institutions in place in order to conduct RRI;
3. Including internal and external stakeholders; and
4. Responding to internal and external stakeholders including CSOs.

In this article, we assume that the four dimensions listed above can be construed as the construct of RRI governance, which we will refer to as the RRI construct hereinafter. This assumption is theoretically grounded in the current state of RRI research [1-3], [5-10].

2.2. The Models of RRI Governance

Another important literature we are building upon is that of models of RRI governance [11-13]. This literature proposes the following four models of RRI governance.

2.2.1. The Standard Model of RRI Governance

The standard model of RRI governance is based on the principle of irrationality. Under this model, the members of a research and innovation consortium assume that other external stakeholders do not have the knowledge and expertise to make sensible decisions regarding the governance of the project. In this model, RRI governance follows a top-down approach that involves only internal stakeholders.

2.2.2. The Revised Standard Model

In the revised standard model, external stakeholders mediate and exert influence on the governance of the project to mitigate potential risks of the project agenda. This leads to an extension of the outside parties involved in the governance of the project, typically involving governmental and regulatory agencies. This leads to higher levels of inclusion, anticipation, reflexivity, and responsiveness than in the standard model.

2.2.3. The Consultative Model

In this model, the internal stakeholders recognize the role external stakeholders beyond governmental and regulatory agencies can play in the governance of the project. This leads to higher levels of inclusion, anticipation, reflexivity, and responsiveness than in the revised standard model.

2.2.4. The Co-Constructive Model

In the co-constructive model, external stakeholders play an active role in the governance of the project, including the co-creation of its outcomes. This leads to the highest levels of inclusion, anticipation, reflexivity, and responsiveness.
3. Research Question and Hypotheses

The central research question we are addressing has been motivated by the interest of research-funding agencies to assess different models of RRI governance.

3.1. The Research Question

How different RRI governance models, expressed in terms of different values for the four dimensions of the RRI construct, might affect research and innovation output and the dynamics of the research funded?

3.2. Working Assumptions

In our model, three working assumptions are postulated, as introduced below.

The RRI Moderating Assumption: The research direction, application domain, and underlying technology focus associated with a project moderate a variable that we will term the “RRI sensitivity” of the project. Projects with high RRI sensitivity are more likely to be affected by norms imposed by agencies providing funding, by governmental bodies and institutions of the private sector, and by CSOs.

The RRI Compliance Assumption: There are overheads associated with the processes of anticipation, reflexivity, inclusion, and responsiveness associated with RRI governance. These overheads increase as we move from the standard model of RRI governance to more inclusive models such as the consultative and the co-constructive models. To the extent that projects are subject to higher levels of RRI sensitivity, there will be an inversely proportional relationship between these overheads, on the one hand, and the innovation output produced by the RI consortium members, on the other.

The CSO Mediation Assumption: CSOs play a mediating function in the governance of RRI by way of exerting power and influence on the agents that are part of project consortia, forcing them to respond to the demands of societal actors and their interest groups.

3.3. Hypotheses

Our model has been designed to test the following hypotheses.

The RRI Trade-Off Hypothesis: Increasing the number of norms associated with the consultative and co-constructive governance models of RRI will have a negative impact on the innovation output achieved.

The CSO Inclusion Hypothesis: Involving CSOs early on in the innovation life cycle will contribute to the success of the project, which we construe as the successful delivery of an innovation output.

4. Theoretical Framework

Considering the four dimensions proposed in [10] to define RRI governance, we can postulate that a project can be characterized as an RRI-compliant project, an RRI project for short, to the extent that the research project has processes in place pertaining to the areas of anticipation, reflexivity, inclusion, and responsiveness. Therefore, we can postulate that the dependent variable responsibility is a function of the independent variables anticipation, reflexivity, inclusion, and responsiveness. All these variables will eventually lead the governing boards of a project consortium to make strategic decisions for which they—and all other members of the consortium—will be held responsible, forcing them to account and respond. Under this model, these dimensions define the “RRI-ness” of the project.
4.1. A Theoretical Foundation of RRI

In the theoretical model shown in Figure 1, white rectangles represent the independent variable \textit{RRI sensitivity}, grey boxes with rounded corners represent mediating variables, grey boxes represent dependent variables, and the white ellipse represents the ultimate dependent variable “RRI project success” of the project.

In our model, the independent variable corresponds to \textit{RRI sensitivity}. Our model is quite a departure from the one suggested by Stilgoe et al. [10]. In fact, we propose that the structure of the RRI construct is more complicated than the one suggested by these authors because in our model the independent variable \textit{RRI sensitivity} influences the dependent variable \textit{responsiveness} of the proposed RRI construct through three mediating variables: \textit{anticipation}, \textit{reflexivity}, and \textit{inclusion} [14]. The dependent variable \textit{responsiveness} will ultimately have an impact on the project success.

4.2. The Problem of Emergence of RRI

The behavior of a project consortium will be influenced by the constraints of an RRI governance model. Such a governance model can be adopted by the consortium as a result of the internal policies and the decisions of the governing board or can be imposed on the consortium by external stakeholders. Consortium members at the micro level might participate in this decision-making process either directly or indirectly, depending on the internal governance model implemented by the consortium. Thus RRI emerges as a result of: (i) interactions of consortium members at the micro level, (ii) interactions of a complex ecosystem of external stakeholders such as CSOs via a process of mediation with the governing bodies of the consortium and the consortium members at the meso level, and (iii) interactions of consortium members and their governing board with the research-funding agencies and regulatory bodies that monitor the activities of the project at the macro level, which can dictate and hence impose macro level constraints of RRI governance upon the consortium members. RRI governance “behavior” will thus emerge in our model as a result of the “concurrent interactions” of a complex ecosystem of actors at the micro, meso, and macro level during simulation rounds.

5. The Agent-Based Model

5.1. The Agents

As in other agent-based models of innovation networks [15-18], we consider the following agents: \textit{instruments}; \textit{calls}; \textit{research-performing agents}; \textit{consortia}; \textit{proposals}; \textit{projects}; and \textit{subprojects}. Research-performing agents are comprised of three types of agents: small and medium-sized enterprises, diversified
firms, and research institutions. The research funding agency is modeled not as an agent but as a single entity that administers instruments, issues calls, and provides funding to agents of type consortium in order for them to instantiate agents of type project and subproject. All of the agent types above have traditionally been considered when modeling the links between universities, industry, and research-funding agencies in innovation networks [15-18]. The new agent in our application domain corresponds to CSOs. CSOs are agents that do not perform research but play an important role in modeling processes of RRI governance in our model.

5.2. The Inputs

Traditionally, agent-based models of innovation networks have endowed research-performing agents with two basic inputs: knowledge stock and capital stock [15-18]. At a higher level of aggregation, agents of type consortium, subproject, and project will also be endowed with knowledge stock and capital stock.

5.2.1. The RRI Construct

In our model, we endow research-performing agents as well as agents of type subproject, project, and CSO as well as regulatory bodies, through the object research-funding agency, with an entirely novel input, namely, the RRI construct. The RRI construct is a tuple containing values for the four dimensions shown in Figure 1.

5.2.2. Research Direction, Technology Focus, and Application Domain

At a meso level of aggregation, we have agents of type proposal. As in previous agent-based models [15], proposals inherit the knowledge stock contributed by research-performing agents at the micro level. We also assume that proposals inherit, as part of the knowledge stock contributed by research performing agents, the research direction (RD) and the technology focus (TF) of all research-performing agents that contributed to the proposal. Agents of type proposal will also have an additional input called application domain (AD). The application domain will be instantiated once the process of proposal writing is complete and the research-performing agents have submitted the proposal to the research-funding agency.

5.3. The Outputs

We borrow from previous work on agent-based models dealing with university and industry relationships [16-18] and with the funding of innovation projects by the European Commission [15]. These models typically include research-performing agents that contribute to producing proposals. If a proposal is successful and funding is provided by the research-funding agency, then agents of type subproject and project will be instantiated. Subprojects are associated with work packages to be conducted by one or more research-performing agents. Research-performing agents that are part of a subproject will produce subdeliverables as outcomes. Such subdeliverables can consist of reports, publications, and unpatented or patented intellectual properties (IPs) that are produced by research-performing agents collectively at the subproject level. The sum of all subdeliverables produced at the subproject level gives rise to the deliverables as outcomes of agents of type project at the meso level. These deliverables are inherited by agents of type consortium and increase the knowledge stock of research-performing agents. Similarly, patented IPs that are produced at the subproject level and correspond to outcomes at the project and consortium level would increase the capital stock of research-performing agents.

5.3.1. RRI Sensitivity

We introduce the RRI sensitivity of a project as an entirely new variable. RRI sensitivity is calculated only if the project proposal is funded by the research-funding agency. In that case, an agent of type subproject or
project is instantiated and inherits the research direction, technology focus, and application domain of the proposal funded. With these three parameters, the RRI sensitivity of the project is calculated. RRI sensitivity of a project is defined as the statistical probability that a project with a given research direction, technology focus, and application domain might come across RRI issues prior to, during, or after its execution. This probability should be assigned based on historical data encompassing all projects with the same research orientation, technology focus, and application domain.

5.3.2. Computing RRI Sensitivity

Let us consider the programs of a research-funding agency. A program may be construed as an instrument, although it would generally be comprised of more than one instrument. An instrument, on the other hand, is comprised of calls. Consortia prepare project proposals in response to calls of the research-funding agency under a given instrument. As per our first working assumption, the technology focus is not per se the determining factor of RRI sensitivity. It is the technology focus in conjunction with the research direction and the application domain that determines RRI sensitivity. Some calls are more prone to leading to projects with a higher RRI sensitive than others. For example, a call with a thematic focus in the area of “ICT for improving the care for elderly people” seems to give rise to project proposals with a lower RRI sensitivity than a call with the thematic focus in the area of “ICT for improving nuclear plant security.”

5.3.3. The Emergence of RRI Sensitivity

RRI sensitivity acts at the project level and is instantiated only for those proposals that are successful in securing funding. Even though this variable acts at the project level, its value emerges from sources at the micro level (the research-performing agents) and the macro level (the research-funding agency). On the one hand, the calls of the research-funding agency give an indication of the technology focus and the potential application domain for project proposals to be eligible. This macro-level information influences the resulting RRI sensitivity of a project. On the other hand, agents that are part of a consortium contribute their knowledge to match the knowledge required to be eligible in a particular call and have some freedom to define the concrete application domain for their proposal. We put the case that the knowledge in a given technology area does not determine per se the RRI sensitivity of a project. In some cases, the knowledge of research-performing agents that is contributed to producing a proposal and will be applied to execute a project might provide some clues regarding the resulting RRI sensitivity of the project. In many other cases, though, the resulting RRI sensitivity will be indeterminate if we look at the knowledge contributed by the research-performing agents alone. As an example, the knowledge that a research-performing agent contributes to a project in the area of public-key cryptography does not necessarily lead to projects with high RRI sensitivity. It will be the application domain, that is, the intended use, the determining factor of RRI sensitivity in such a case. In other cases, the technology focus alone might suffice to render the project highly RRI sensitive, as in the case of projects involving the use of nuclear energy regardless of its application domain.

5.3.4. RRI Sensitivity as a Moderating Factor of RRI

The RRI sensitivity acts as a “moderating” variable, that is, it creates the conditions for a project to allocate more resources in each of the four dimensions of the RRI construct in order for the project to be labeled “RRI compliant.” Therefore, a project with a low RRI sensitivity will require the dimensions of its RRI construct to adopt lower values in order to render the project RRI compliant.
5.3.5. Social Stock

We endow research-performing agents as well as agents of type consortium, project, subproject, and CSO as well as any regulatory bodies with social stock. Social stock results from the values of the dimensions of the RRI construct of these agents. Our motivation to endow these agents with social stock is motivated by Nahapiet and Ghoshak [19]. Of particular importance for our present purposes is the work of Coleman [20], who distinguishes trustworthiness, information sharing and relational norms and sanctions as the three dimensions of social capital. These three dimensions are closely connected with the four dimensions of the RRI construct.

5.4. The Model

In this section, we introduce the agent-based model.

5.4.1. Agent-Based Model Definition

Definition 1 (Agent-Based Models) Let a group of agent-based models be the tuple $\Theta = (\mathbf{T}, \mathbf{I}, \mathbf{G}, \mathbf{A}, \mathbf{O})$, where $\mathbf{T}$ is a set of different types of agents, $\mathbf{I}$ is the set of inputs, $\mathbf{G}$ is the set of decision gates, $\mathbf{A}$ is the set of actions, and $\mathbf{O}$ is the set of outcomes.

Definition 2 (Agent-Based Model) Let an agent-based model be the tuple $\theta = (\mathbf{I}, \mathbf{G}, \mathbf{A}, \mathbf{O})$, where $\mathbf{I}$ is the set of inputs, $\mathbf{G}$ is the set of decision gates in an agent-based model, $\mathbf{A}$ is the set of actions, and $\mathbf{O}$ is the set of outcomes. For each agent type, there will be a tuple associated with its agent-based model.

5.4.2. Agent-Based Decisions as a Multicriteria Decision Problem

This decision-making process of agents in our agent-based model builds upon the evaluation process in [21].

Definition 3 (The Agent-Based Decision Problem) Let the agent-based decision problem $\tau$ be defined as the triplet $\tau = (\mathbf{G}, \mathbf{P}, \Pi)$, where $\mathbf{G}$ is a set of decision gates, $\mathbf{P}$ is a set of points of view, and $\Pi$ is a multicriteria decision problem.

Definition 4 (The Agent Evaluation Problem) Let the evaluation problem be the tuple $\rho = (\mathbf{A}, \mathbf{C}, \mathbf{U}, \mathbf{L})$, where $\mathbf{A}$ is set of alternatives derived from $\mathbf{G}$, $\mathbf{C}$ is a set of criteria derived from $\mathbf{P}$ allowing the evaluation of elements of $\mathbf{A}$ using the criteria in $\mathbf{C}$, $\mathbf{U}$ is a model of the uncertainty regarding the information in $\mathbf{A} \times \mathbf{C}$, and $\mathbf{L}$ is an aggregation logic defining how the information in $\mathbf{A}$ and $\mathbf{C}$ will be used to obtain a solution to the problem $\Pi$.

Definition 5 (The Recommendation) An alternative in $\mathbf{A}$ is chosen based upon the recommendation that results from the aggregation logic $\mathbf{L}$ and the evaluation criteria in $\mathbf{C}$.

5.4.3. The Dynamics of the Model

The flow chart shown in Figure 2 describes the dynamics of the agent-based model. In Figure 2, the white rectangle with rounded corners represents the initialization module, white rectangles represent inputs, grey rectangles represent outputs, white diamonds represent decision gates, circles represent decision gates with two possible outcomes (pass and fail), white ellipses represent alternatives, grey ellipses and circles represent objects, and grey rectangles with rounded corners represent actions. The first part of this agent-based model is based on the dynamics of other agent-based models based on SKIN [15-18]. This part of the model encompasses an initial stage comprised of three main activities: consortium formation, proposal writing, and proposal evaluation. The second stage is the most relevant one in terms of simulating processes of RRI governance. The main activities of this second stage are: RRI sensitivity computation, RRI construct computation, deliverables creation, and CSO mediation. CSO mediation plays a central role in modeling the process of CSO intervention during processes of RRI governance.
5.4.4. Simulation of RRI Governance Policies

Our model has been designed to simulate policy intervention in the area of RRI governance and the role of CSOs as mediators of RRI governance. Not only the impact of policies of RRI governance on the actions consortia formation, proposal creation, and proposal evaluation should be simulated but, more importantly, their impact on project success or failure in terms of level of project completion and overall project “RRI-ness.” This is accomplished by setting up in the model a so-called “regulatory RRI construct” with values that can be imposed on consortia by regulatory bodies. The values of this construct can be adjusted by the terms of a call issued by the research-funding agency (FA in Figure 2). In the absence of any CSO intervention, the regulatory RRI construct as well as the RRI construct of agents (RRI construct in Figure 2) and the RRI sensitivity of the project are inputs to the action called RRI construct computation. This action is supervised by the governing board. Adjustments to the RRI construct of a project are made via the action RRI construct computation process using a multicriteria decision analysis algorithm described in [22]. As we will see, the RRI construct of CSOs can also be an input to the RRI construct computation process.

5.4.5. CSO Mediation

We also model the impact of CSO intervention on the governance of RRI. To this end, we have introduced CSOs as agents in our model. CSO-involving events occur once a subproject has been instantiated and starts producing subdeliverables. The occurrence of these CSO-involving events is modeled using a Poisson
probability distribution. A random CSO-involving event, which will be very rare for a project with a low RRI sensitivity, will become more likely as the RRI sensitivity of the project increases. During a simulation, a project might need to deal with more than one CSO. This will be the case when a number of CSOs begin to “intervene” with the governance of the project. CSOs have an RRI construct and they also have an importance. The CSO mediation process will then involve the RRI construct of the subproject at the agent level, the RRI constructs of all interfering CSOs, and the regulatory RRI construct. For scalability reasons, we assume that only CSOs over a given “threshold” of importance will be considered. These inputs are then made available to the governing board in order for it to supervise the process called RRI construct computation. The board will then decide the level of adjustments they are willing and able to make to the values of each of the variables in the subproject RRI construct via a process that we call CSO mediation in Figure 2. We also model this process as a multicriteria decision process. We refer the reader to [22] for a detailed description of the multicriteria decision analysis method used to implement the process of CSO mediation in our agent-based model.

6. Conclusions

In this article, we have put forth the first agent-based model of RRI governance. Governance of RRI, such as the modeling of other phenomena in the social sciences, is an area where the interactions of variables at the micro, meso, and macro levels “concurrently” influence the way social processes emerge in the real world. We have put forth our model with this in mind in order to avoid the fundamental flaw of agent-based modeling approaches that try to “force” the dynamics of their models in such a way as to emerge as a result of decisions made solely by agents at the micro level. These approaches assume, incorrectly, that the behavior of complex social systems emerges out of decisions made by agents at the micro level in isolation, that is, disregarding the constraints imposed by other entities at the meso and macro levels. As a result, these models are difficult to reconcile with the way processes of innovation and other social phenomena unfold in the real world and are generally at odds with the way behavior emerges in complex systems in the social sciences [23]. The governance of RRI [1-3], [5-10] provides us with plenty of examples of how the concurrent interaction of features at the micro, meso, and macro levels influences the emergence of behavior in complex social systems.

Our work has also been inspired by the challenges innovation managers face when modeling the complex trade-offs posed by CSOs intervening in research and innovation agendas. The complexity of these trade-offs calls for the utilization of more complex methods for decision making. Our agent-based model is not only the first to deal with RRI but it is also the first to include multicriteria decision analysis methods to deal with these complex trade-offs. We refer the interested reader to [24] for a description of how to extend agent-based modeling through multicriteria decision analysis in order to simulate more complex social phenomena.

Future work will consist in utilizing multicriteria decision analysis methods that deal with hybrid and fuzzy data for a characterization of the four models of RRI governance described in this article [25-26]. We will also model the impact that more democratic models of RRI governance may have on innovation output. To this end, we will investigate the impact of new governance models based on the principle of bottom-up governance. Bottom-up models of RRI governance have not yet been proposed and are virtually not in operation in the area of “academic” RRI. Such governance models would transfer the RRI governing powers from governing boards to committees comprised of internal and external stakeholders. Members in such committees could nominate representatives and delegate RRI-governing powers or they could assume a more active role in the governance of RRI through processes of direct collective action [27].

Acknowledgements

The first author has been partially supported by the GREAT project, a project in the area of Governance for Responsible Innovation funded by the European Commission. The agent-based model of RRI presented in this article departs from the view on agent-based modeling of RRI currently being adopted in the GREAT project.
Therefore, the model yet to be developed in the GREAT project does not necessarily endorse any of the ideas presented in this article.

References


