Ethics and Information Technology (2024) 26:22
https://doi.org/10.1007/s10676-024-09753-x

ORIGINAL PAPER

# Tailoring responsible research and innovation to the translational context: the case of AI-supported exergaming

Sabrina Blank[1] · Celeste Mason[2] · Frank Steinicke[2] · Christian Herzog[1]

Accepted: 8 February 2024


## Abstract

We discuss the implementation of Responsible Research and Innovation (RRI) within a project for the development of an AI-supported exergame for assisted movement training, outline outcomes and reflect on methodological opportunities and limitations. We adopted the responsibility-by-design (RbD) standard (CEN CWA 17796:2021) supplemented by methods for collaborative, ethical reflection to foster and support a shift towards a culture of trustworthiness inherent to the entire development process. An embedded ethicist organised the procedure to instantiate a collaborative learning effort and implement RRI in a translational context. Within the interdisciplinary setting of the collaboration and with the support of a technoethicist, we successfully identified relevant, project-specific challenges and developed a roadmap with derived actions, thus meaningfully integrating RRI into the development process. We discuss the methodological procedure in terms of its effectiveness and efficiency, the allocation of responsibilities and roles, particularly regarding potential frictions in the interdisciplinary context with embedded ethics, and the challenges of the translational context. We conclude that the responsibility-by-design standard effectively established a productive workflow for collaborative investigation and work on ethical challenges. We reflect on methodological difficulties and propose possible avenues to our approach.

**Keywords** Responsible research and innovation · Industry · Trustworthiness · Medical artificial intelligence · Interdisciplinary collaboration

## Introduction

The persistent gap in the transference of AI ethics principles into technical development processes is a common and recurring issue in recent discussions (Hallensleben et al., 2020; Morley et al., 2020; Petersen et al., 2022). Even when aware of potential ethical implications of the innovation, a project team developing technological innovations usually lacks the resources, capacity, or knowledge to identify *what* ethical risks are relevant, *when* to address them in the development process, and *how* they can be mitigated (Breyer & Herzog, 2022). As research projects are usually envisaged with scarce resources (budget, time, capacities), leading to a lack of flexibility and responsiveness (Morrison et al., 2020), there is a need for lean and efficient processes for socio-ethical reflection and anticipation, while at the same time ensuring the necessary degree of ethical depth and mitigation. The collaboration's objective was to *identify*, *analyse* and—if possible—*address* socio-ethical issues. Therefore, a development team for an AI-supported exergame, including an embedded ethicist and an external technoethicist, collaborated to explore and examine methods to help (1) reflect on and anticipate the ethical implications of the innovation and (2) strategically align the ethical issues with the development process, including deriving options for action. The embedded ethicists' additional objective was to explore *how* such a procedure can be conducted in translational research—a collaboration between academia and industry to turn research results into products—with minimal resources.

✉ Sabrina Blank
sabrina.blank@uni-luebeck.de

Celeste Mason
celeste.mason@uni-hamburg.de

Frank Steinicke
frank.steinicke@uni-hamburg.de

Christian Herzog
christian.herzog@uni-luebeck.de

[1] Ethical Innovation Hub, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany

[2] Department of Informatics, University of Hamburg, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany

Published online: 26 March 2024

This approach aims to contribute to the translation of RRI into industry practice, focusing not only on the ethics of the product but also on an inclusive and transparent development process that is responsive to the values of the target group and societal and moral norms. We adopted the responsibility-by-design (RbD) approach (CEN CWA 17796:2021; Porcari et al., 2019), supplemented by the Ethics Canvas method (Reijers et al., 2018), and adapted the procedure to the needs of the project.

We commence briefly reviewing select issues in RRI. This includes various views on the (non-)embedded role of ethicists in technoscience and frictions in interdisciplinarity, the translational setting of academic-industry collaboration, and challenges and barriers of implementing RRI in industry. We then outline the project's content and collaboration's requirements. We illustrate our methodological procedure, describe and discuss the selected methods. We look into processing workshop outcomes in line with identifying, analysing, and addressing possible socio-ethical implications. In addition, we sketch the workshop outcomes to demonstrate the achievement of the collaboration. Finally, we discuss the methodological approach—its effectiveness and efficiency, questions of responsibility allocation and consensus within interdisciplinary teams, and the challenges of the translational research context. The discussion also reflects on the role of embedded ethics and potential frictions. We conclude that using RbD, augmented by further collaborative modes to foster interdisciplinary ethical reflection, can effectively devise project-specific and actionable project plans that consider the financial and human resources required to move from translational research to trustworthy development processes for AI-supported systems.

## Challenges and barriers of and with RRI and the need for translational research

The discussion on RRI ranges from the underlying fundamental principles (Burget et al., 2017; von Schomberg, 2013), its embedding and contextualisation (Demers-Payette et al., 2016; Pacifico Silva et al., 2018), up to the design of strategic and operative concepts and standards for implementing RRI in industrial practice (van de Poel et al., 2017; CEN CWA 17796:2021; Porcari et al., 2019). In particular, the challenges and conceptual underpinnings of implementing RRI in industry are the focus of recent debate (Martinuzzi et al., 2018; Ribeiro et al., 2018; Yaghmaei & Poel, 2020; Blok et al., 2020; Gurzawska et al., 2017; Dreyer et al., 2017). Our research aims to contribute to the advancement of practical methods for translating RRI into practice, with a specific focus on AI-based medical innovations.

Scholars have raised major criticisms towards the conceptualisation and implementation of RRI, especially in industry and translational research settings. Nazarko (2020) criticises RRI's Western Eurocentrism, which overlooks cultural and regulatory differences between countries that matter for industrial actors. Additionally, the demand for transparency, often raised in RRI contexts, conflicts with protecting intellectual property and advantages in competitive economic environments. Owen et al., (2012) bemoan RRI's inherent idea of collective responsibility, which supposedly leads to unclear expectations of the role of the involved actors. For instance, the role of ethics is widely debated between two extreme ideals: integration and embeddedness versus autonomy and critical distance. McLennan et al. (2022) distinguish levels of embeddedness for ethicists, suggesting that the "gold standard" would have ethicists as team members fully integrated into the development process. Embedded ethicists can support responsible development by ensuring the societal value of innovations, analysing risks and harms, and anticipating long-term effects on the socio-technical ecosystem and society. Thus, embedded ethicists should assist in proactively considering ethics in anticipation rather than reactively responding to social and ethical challenges alone (McLennan et al., 2022).

Regarding this matter, Zwart et al. (2014) argue that embedded ethics encounter a dilemma between their proximity to and their autonomy from the project. In such situations, the critical distance essential to ethical assessment might be compromised by the need to serve the project's aims. Consequently, ethicists may become "pre-formatted" and "co-responsible" with the project's objectives, undermining unbiased ethical evaluation. Further tensions between disciplines arise due to hidden agendas, economic pressures, limited resources, and competition (Kuzma & Roberts, 2018). Rommetveit et al. (2019) identify a significant barrier in weighing qualitative against quantitative evidence due to the predominance of qualitative approaches in ethics and the reliance on quantitative approaches in technoscience. However, additional efforts in interdisciplinary collaborations often conflict with the need for rapid technological development, which, in turn, may impede thorough ethical considerations. This conflict can arise due to limited funding or market dynamics, which may demand shortcut ethical input to avoid slowing down processes (Kuzma & Roberts, 2018; Rommetveit et al., 2019). However, interdisciplinary work requires time for communication (McLennan et al., 2022). This poses a risk of implementing one-size-fits-all approaches, which may obscure project-specific ethical implications (Rommetveit et al., 2019). As a result, interdisciplinary work faces challenges in combining disciplines on a process level and suffers from a lack of competencies on both sides—technoscientists for ethics and ethicists for technosciences—to effectively integrate their expertise (McLennan et al., 2022). Rommetveit et al. (2019) advocate that "interdisciplinarity is an outcome to be achieved, and

not a default starting point". However, Kuzma and Roberts (2018) suggest that in the short term, RRI scholars should focus on creating supportive tools and concepts for specific innovation systems contexts rather than demanding a complete sustainable system change to avoid overburdening and potential resentment from innovators and regulators. Kuzma and Roberts (2018) state that collaborative support is crucial in navigating the appropriate and realistic means of compromise. It is important to note that this does not undermine efforts towards societal alignment of long-term innovative solutions.

In summary, there are varied opinions regarding the role of ethicists, ranging from altering the societal value of innovation by being involved and embedded in the process, diagnosing policies and integrating with policy processes to remaining outside the industry to avoid dependence on the project goals (Rommetveit et al., 2019). The discussion regarding extending responsibility (Grunwald, 2011; von Schomberg, 2013) from individualistic to interdisciplinary notions (Stahl et al., 2021) informs our view of a collaborative and participatory approach. This will be explored further in the discussion section, considering the interdisciplinary approach of this case study.

In the context of AI-based technology and the RRI agenda, an almost dominant discussion currently concerns the so-called principle-to-practice gap (Mittelstadt, 2019; Morley et al., 2020; Whittlestone et al., 2019). The gap refers to an alleged lack of methods, tools, and knowledge to rigorously apply principles—i.e., any set of requirements, normative ethics or RRI, however broadly conceived—to development practice concerning the product and the process. In terms of practical impact, trustworthiness in this sense requires that innovation processes are aligned with RRI ideals *and* can address the specific implications revealed by AI ethics, such as algorithmic bias (e.g., Koene, 2017; Wachter et al., 2020), explicability (e.g., Floridi et al., 2018; Herzog, 2021), etc. Accordingly, it must be acknowledged that realising RRI in a translational research or industrial context is difficult.

Our research is therefore oriented towards the general issue that "RRI links up with two key challenges […]: the accelerating global race to innovate in order to maintain competitive advantage, and the struggle to maintain public trust in business" (Martinuzzi et al., 2018). Martinuzzi et al. (2018) raise several challenges from an industry perspective. These include a lack of awareness of the concept of RRI, uncertainty about the benefits, and RRI methods being perceived as a barrier rather than a competitive advantage (Martinuzzi et al., 2018; Ribeiro et al., 2018). Blok et al. (2020) and Dreyer et al. (2017) describe further challenges: an ambiguity of potential impacts as the technologies are at an early stage of development, intellectual property issues in light of participative and inclusive design approaches, a lack of consumer awareness, and barriers derived from the institutional environment. Another prominent challenge is that the diversity of approaches does not adequately refer to business-oriented tools to support implementation, as RRI has mainly been used in publicly funded projects (Martinuzzi et al., 2018) and thus fails to generate governance impacts (Dreyer et al., 2017; Ribeiro et al., 2018). This highlights the need to create incentives for industry in RRI, which have been elaborated in terms of business incentives ranging from external demands from critical consumers, corporate reputation, and related certifications to the benefits of engaging employees as internal stakeholders, as well as boosting the sustainability and societal desirability of innovation (Gurzawska et al., 2017). Consequently, some requirements must be met to implement RRI in industry successfully: Frameworks must be better aligned with established industry practices, and RRI needs to help clarify its purpose regarding the added value (Dreyer et al., 2017).

We consider the translational setting, where academia and industry collaborate on the implementation of state-of-the-art AI methods in the healthcare sector, to be a fruitful area for answering the question of what is needed to put RRI methods into practice for at least three reasons:

(i) Translational research can be a point of contact between RRI science and industry to pilot RRI practice with reduced economic risks (Porcari et al., 2019).
(ii) Translational research can incentivise RRI in science and industry to assess their cooperation, producing meaningful outcomes.
(iii) Translational research reinforces the importance of RRI practices, such as ethical foresight analysis (Floridi & Strait, 2020), to plan and mitigate risks.

We identified the RbD standard (CEN CWA 17796:2021) as the most promising candidate framework for our project (see detailed reasoning in the section on methodological procedure). As RbD focuses on the company level and requires an extensive process with significant commitment, several open questions remain relevant to the context of *translational research* at the *project level*:

- Is it possible to facilitate the development of a shared RRI vision for the innovation supported by the innovator's leadership, even though the latter might not be directly involved?
- What are the resource constraints and participants' competencies in planning for future RRI actions? Can necessary resource allocations be identified and secured?
- Can concrete and product-specific challenges, barriers, drivers, and risks be derived? Can these outcomes be

transferred to other innovation processes/product innovations?

The context of translational research gives rise to the following vital questions:

- Will the skills and competencies of the project team from the translational context be transferred and available in the product development phase after the academic-industry collaboration has ended?
- Will the collaborative processes in translational research sufficiently mimic industrial innovation processes so that the lessons learned and methods used in the RRI domain will continue to be adopted when development enters the next phase, with much less input from academia and potentially quite different workflows?

Because of these open questions alone, RRI case studies at the translational project level are an attractive way to demonstrate the methodological approach of strategic integration of RRI, which needs to be critically analysed and reflected upon after the fact. Projects have different circumstances and requirements, making a project-specific procedure crucial. Our paper aims to inspire to promote the development of best practices further.

## The project and requirements of the collaboration

This collaboration is part of an innovation ecosystem called KI-SIGS[1]. The objective of KI-SIGS is to develop an implementation concept for an "AI Space" platform for intelligent health systems. Several projects and stakeholders, such as companies, hospitals, and universities, engage in this ecosystem. Within this innovation ecosystem, the so-called Responsible Innovation Platform is dedicated to establishing guidance for responsible innovation in medical AI-based technology. This article focuses on the interdisciplinary collaboration between this platform and the project team for an AI-supported exergame. The former includes an embedded ethicist who is an early career technoethicist with expertise in business ethics and practical philosophy. Throughout this collaboration, in addition to providing ethical expertise, the embedded ethicist was also responsible for tailoring the RbD process to the constraints and requirements of the project preparing and moderating the workshops. It also includes an external technoethicist with a background in engineering and applied ethics, providing additional ethical expertise in technological innovation at some stages. The latter comprises engineers, informatics, and content developers in human–computer interaction, mixed reality, and exergaming. The collaboration aims to promote responsible innovation in the project's development and implementation.

In the project, partners from the Hamburg University[2], the University of Bremen[3] and the companies SZENARIS[4] and apoQlar[5] are developing an AI-supported exergame to improve the health and fitness of older adults with the support of the Hospital zum Heiligen Geist,[6] Hamburg. This project addresses how to motivate older adults to exercise more often and more effectively and how to support them in integrating physical activity habits more easily into their daily lives. To develop an efficient training system, the design followed a user-centred design process to provide a feature-rich system that is tailored to the needs of the target users—primarily older adults aged 55–67, who need help to develop and maintain health, mobility, and fitness habits without the constant supervision of trainers, physiotherapists, or medical staff.[7] Towards this goal, the application integrates gamification elements with intelligent virtual and robotic agents into motion training systems that can provide users with specific feedback on their movements using motion detection algorithms.

The collaboration's goal was to develop a strategic plan to identify ethical challenges and allocate future RRI activities while simultaneously providing a bridge to operational activities that promote the marketability of the innovation in a responsible and socially desirable way. However, the project entailed requirements to be considered when planning the collaboration, particularly in selecting appropriate methodologies.

- The project plan did not consider resources required for collaboration.
- Limited resources of the platform due to other commitments with other projects in the innovation ecosystem
- A need for virtual/distance collaboration due to the ongoing coronavirus pandemic and non-collocation of project partners

Thus, the modes of collaboration had to be tailored to balance lean processes and methods enough to achieve significant depth in the collaborative, ethical reflection and discussion. Given the amount of technical development and

[1] AI spaces for intelligent health systems, engl. translation for the German "KI-Spaces für intelligente Gesundheitssysteme"

[2] Universität Hamburg, Mittelweg 177, 20148 Hamburg, Germany

[3] Universität Bremen, Bibliothekstraße 1, 28359 Bremen, Germany

[4] SZENARIS GmbH, Otto-Lilienthal-Str. 1, 28199 Bremen, Germany

[5] apoQlar GmbH, Raboisen 32, 20095 Hamburg, Germany

[6] Hospital zum Heiligen Geist, Hinsbleek 11, 22391 Hamburg, Germany

[7] In the following, we will refer to physiotherapists

translational research required to get close to the product level, a two-pronged strategy was adopted regarding ethical and responsible innovation issues: Address the most immediate issues immediately but plan and consider the additional resources required to properly investigate issues over a longer time horizon.

As discussed below, the translational research setting remained a guiding consideration in designing individual workshop sessions within the collaboration.

## Methodological procedure and outcomes

In this section, we describe our methodological procedure and why we decided to use particular methods. We also provide a glimpse of the outcome of the ethical analyses, which helps to illustrate our reasoning or supports observations made in the following discussion. However, we limit the presentation of outcomes from the RRI project because this paper focuses on our methodological approach to facilitating the transfer of RRI into practice, with AI-supported exergaming serving only as a recent case study. Therefore, when we talk about results, we will refer to the results of our RRI procedure rather than the actual outcome of, for example, ethical analyses.

We adopted the RbD methodology, which has been assessed through nine company case studies in the EU PRISMA project (Porcari et al., 2019), by adapting the six steps for the RRI roadmap design to our project needs. While in the PRISMA project, all activities were conducted at company level, we focused on a single translational research and innovation project, with responsibilities distributed among different partners from industry and academia. This means that we have adapted the design to one innovation development process and its team rather than to the executive level of a company. The steps related to our variation of the approach are described below.

Step 1: Top management commitment and leadership endorsement towards RRI was obtained after several meetings focused on mutual familiarisation and organisation

Step 2: Context analysis for the identification of ethical, legal, and social impacts was carried out in a workshop using the Ethics Canvas as supportive method

Step 3: Materiality assessment to identify and prioritise drivers and challenges, risks and barriers, stakeholders, and actions was facilitated in two separate workshops

Step 4: Experiment and engage with pilot actions and stakeholders has not yet been carried out

Step 5: Validate impact of the roadmap has not yet been carried out

Step 6: Roadmap design was created as a draft for further evaluation

Steps 4 and 5 still need to be carried out. Accordingly, the resulting RRI roadmap explicitly considers the resources needed to experiment with pilot RRI actions and engage with stakeholders to further inform future aspects of the RRI roadmap. We have postponed validating the roadmap's impact on the product until the product is closer to market.

The reason for selecting the RbD standard was that it considers several business standards and norms (CEN, ISO) and incorporates the method of an RRI-CSR roadmap to integrate RRI strategically along the research and innovation process. Using the RRI roadmap, we aimed to create a strategic integration of ethical considerations in the project, directly aligned with the development process, to plan what, when and how to prevent or mitigate. This served the collaboration's goal of developing a strategic plan and providing a bridge to operational activities. For the context analysis, we chose the Ethics Canvas as a supportive method because it does not presuppose ethical expertise like other tools (e.g., value-sensitive design, ethical impact assessment) and encourages collaboration (Reijers et al., 2018). It provides a low-threshold entry point into collaborative, ethical analysis and stimulates communication within the project team as it is loosely guided by traversing given thematic fields. This supported our goal of facilitating the project team's competencies in reflecting and anticipating the ethical implications of the project's ecosystem. In addition, the methodological similarity to the Business Model Canvas (Business Model Foundry AG, 2014), and thus a presumed immediate familiarity for the project partners from industry, is intended to provide a lean and efficient process. This meets one of the core needs of industry to implement RRI, as mentioned above.

### The procedure in a nutshell

Our procedure of interdisciplinary collaboration follows a participatory approach involving the whole project team, an embedded ethicist, and an external technoethicist at specific points in the process. We commenced with a workshop using the Ethics Canvas for preliminary identification of ethical implications and context analysis. In a second workshop, we started to prepare the RRI roadmap to create an RRI vision and to map key aspects of the development process. This served as a structural link to a third workshop, where the outcomes from the Ethics Canvas workshop were integrated into the roadmap and related actions were derived in line with the development process. The embedded ethicists used the time between the workshops to process the outcomes further and prepare for the next workshop. Figure 1 displays the main procedural steps, their outcomes, the participants involved, and the most significant research result from the reflection of each procedural step.

In the subsequent sections, we present the rationale and design of the Ethics Canvas, the RRI roadmap workshops, and intermediate steps in more detail.

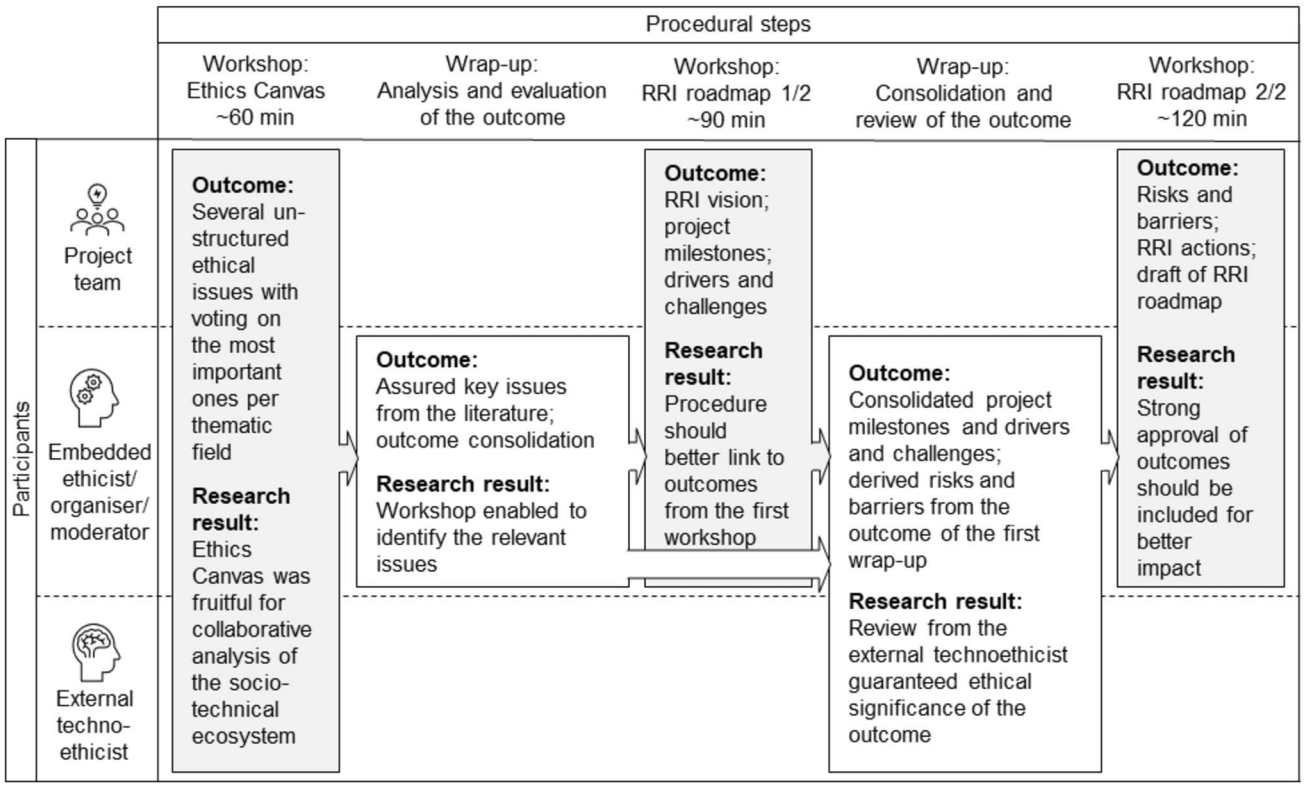


**Fig. 1** Main procedural steps, their outcomes, the participants involved, and the most significant research result from the reflection of each procedural step

## Workshop: Ethics Canvas

After a kick-off meeting aimed at commitment and coordination, the collaboration commenced with an online workshop using the Ethics Canvas (Reijers et al., 2018) to initiate collaborative reflection and discussion on the ethical implications of the project. The aim of the first workshop was a preliminary context analysis of relevant ethical issues within the socio-technical ecosystem of the project. The Ethics Canvas is structured into nine thematic fields, ranging from *affected individuals* and *groups* and their *changing behaviours, relations* and *worldviews,* as well as related *emerging conflicts* over *product failure* and *problematic use of resources,* to the thematic field of action in terms of *what to do* about the identified aspects. Thus, it triggers different perspectives for anticipation and reflection. Regarding the limitation of time and capacity, integrating the Ethics Canvas's shift from individual to group level between the subsequent thematic fields allowed for a more streamlined process, so we worked on six instead of nine thematic fields in the workshop (see Appendix Table 1). The last thematic field of *what to do* was only used to get an idea of what to prepare next.

The entire project team, an embedded ethicist who simultaneously functioned as organiser and moderator, and an external technoethicist were involved. The workshop procedure consisted of three phases with a total of 60 min. The workshop was divided into three steps:

1. Each participant takes notes about the aspects of the thematic fields to ensure that each participant's perspective is presented in the brief time available, rather than having a direct open discussion that risks undermining some contributions (20 min).
2. One starts to present its notes from the first thematic field, and the others iteratively add only new aspects so that each perspective is depicted, but without repetitions (25 min).
3. Each participant votes for the two most critical issues of each thematic field to get a first approximation of the most relevant topics, which we will then briefly discuss in plenary (15 min).

The workshop outcomes are sketched in Table 1 in the appendix to give a brief overview of the issues identified and provide a rough indication to better retrace further processing. The table shows a streamlined version of the outcomes in terms of consistent wording, the same abstraction level, and a rough structure of the voting per thematic field.

Accordingly, the first workshop produced several unstructured outcomes regarding the ethical aspects of context

analysis without clearly prioritising what to justifiably focus on. Therefore, the embedded ethicist performed a wrap-up to consolidate and compare the outcomes with the literature. Only then did we consider the next step, in which possible actions were derived.

### Wrap-up of the workshop: Ethics Canvas

After the workshop, a short presentation of the main outcomes was given in the regular project team meeting to discuss and confirm the outcome. The wrap-up's aim was to analyse the outcomes and the assurance of known key issues from literature in terms of AI-supported exergaming. The analysis and evaluation considered a triangulated approach: the outcomes and voting of the project team, the comparison with findings from the literature, and the estimation of the embedded ethicist.

The outcomes of the workshop were first analysed within each thematic field. The voting was a tool to create a quick but preliminary indicator of the project team's perceptions, which served as a helpful starting point for each thematic field to look at its higher-voted aspects. The structuring allowed for a better overview of the issues identified. The revised outcomes were then roughly compared with any consistent mention of prominent issues in the literature to ensure the relevant aspects were in view. The literature analysis focused on ethical findings concerning exergames and AI issues, as these are the project's main topics. The main implications of the comparison are presented below.

These relate to psychosocial effects of exergaming, such as social isolation and loneliness (Bhattacharya et al., 2022; Li et al., 2018; Benzing & Schmidt, 2018), obstacles in learning how to use and maintain the product (Bhattacharya et al., 2022), stress or frustration from competitive or cooperative modes and insufficient integration into daily life (Rüth & Kaspar, 2021), as well as the possibility of unhealthy overuse or addiction (Bhattacharya et al., 2022; Benzing & Schmidt, 2018), etc., which could be induced. In addition, the consequences of ineffective training or harm (Tobaigy et al., 2018) are related to technical reliability on the one hand and the actual individual use of users on the other. In this context, the need for adequate instructions for use supervised by physiotherapists, could create pressure by requiring technical competence and a shift in their role (Tobaigy et al., 2018). The rough comparison with the literature showed that at least the most critical aspects of exergames were anticipated.

The regrouping and processing of the outcomes allowed a clear preparation and specific elaboration of the ethical focus issues through the expertise of the embedded ethicist. It was more of an evaluation of which issues should be worked on further and which should not be missed at this stage rather than a decision on priorities because of a lack of assessment of the interests of external stakeholders and the actual feasibility of preventing or mitigating certain issues. In this way, the wrap-up produced more structured and better-evaluated outcomes through consultation with the workshop participants, the literature and the embedded ethicist's expertise.

The wrap-up also included the preparation and organisation of the following second workshop.

### Workshop: RRI roadmap 1/2

In this workshop, we started to prepare the RRI roadmap. The RbD standard's RRI roadmap methodology suggests first formulating an *RRI vision* to ensure the RRI commitment and product alignment. The roadmap itself is designed with the abscissa (x-axis) mapping the time to market and the ordinate considering the *drivers and challenges* to realise the project responsibly, the *risks and barriers* to be addressed by relevant *RRI actions* in line with the *milestones* to reach the market (CEN CWA 17796:2021; Porcari et al., 2019) (see Fig. 2).

In consultation with the project team, it was agreed that the RRI roadmap would be developed in two online workshops to better fit the schedule and avoid losing focus. This workshop aimed to create an RRI vision, identify project milestones and formulate project drivers and challenges. The project team attended the workshop, and the embedded ethicist acted mainly as moderator, only actively participating in ethical discussions. The workshop lasted 90 min and was organised as follows:

- Creation of the RRI vision, starting with each participant of the project team formulating their vision in 1–2 sentences to again include each perspective, followed by reading out each individual vision and voting for the personal best version; then discussing the outcomes and compiling the final version of the project's RRI vision (30 min)
- Identification of project milestones as well as drivers and challenges, divided into two groups working simultaneously, as in this step, the issues are general project issues, so we refrained from explicitly gathering each perspective (30 min)
- Presentation and discussion of the outcomes with a decision on the key aspects to be included in the roadmap (30 min)

The elaborated RRI vision focuses on the well-being of older adults by increasing motivation for physical activity through exergames (see Fig. 2). A first outcome regarding the product milestones is a categorisation of the collected aspects into user-oriented, technically-oriented and safety/support issues. The drivers and challenges cover societal, user-related, technological, and economic issues. The key

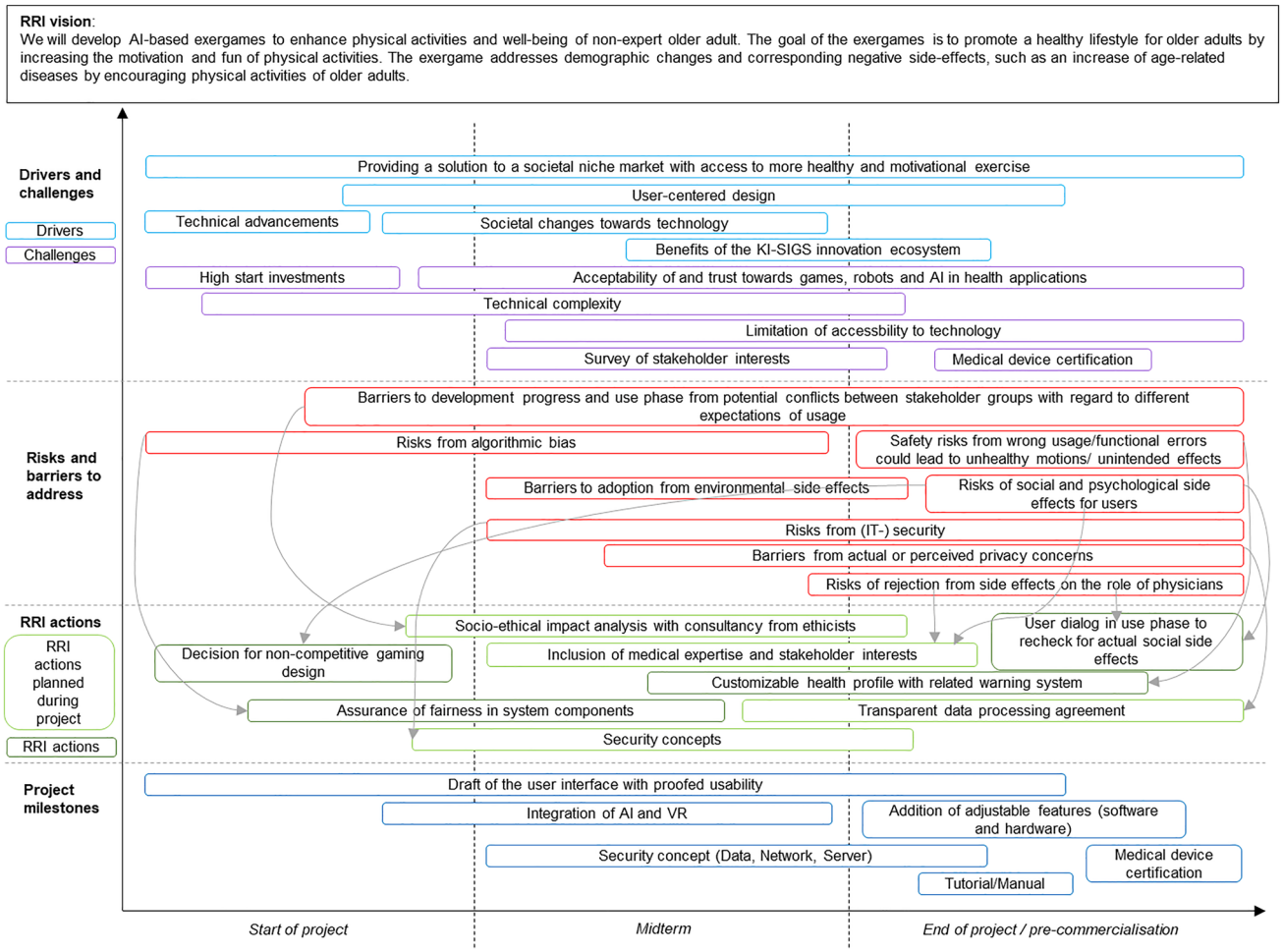


**Fig. 2** RRI roadmap of the project (own illustration inspired by CWA 17796:2021)

finding is that the opportunity for innovation could become a pitfall if it is not adequately developed, integrated into the health system, and used. In particular, acceptance of and trust in the innovation and the inclusion of stakeholder interests are important drivers but also challenges of innovation.

## Wrap-up of the workshop: RRI roadmap 1/2

This wrap-up aimed to ensure the clarity of the roadmap by providing the necessary consolidation of the outcomes. This was achieved by merging related topics and structuring along the time-to-market axis as a proposal that can enter a state of discussion. Looking ahead to the risks and barriers outlined in the roadmap design, the main themes from the first wrap-up on the outcomes of the Ethics Canvas were transformed into risks and barriers to aligning them with the more abstract level of drivers and challenges. This ensured the manageability of the roadmap and led to a further summary of issues that could be merged. The additional external technoethicist reviewed this procedure to ensure ethical significance and comprehensibility. In the following workshop, these outcomes had to be discussed and confirmed to be used as a basis for deriving actions.

## Workshop: RRI roadmap 2/2

The objective of the second workshop on the RRI roadmap was to confirm the revised milestones, drivers and challenges from the first workshop and to discuss the risks and barriers extracted from the Ethics Canvas. Furthermore, we elaborated on actions to address the risks and barriers. The project team and the embedded ethicist participated in this workshop. The workshop lasted 120 min and was structured as follows:

- Agreement on the revised milestones, drivers and challenges to ensure the relevance and appropriateness of the issues finally mapped in the roadmap (15 min)

- Presentation of the further processed outcomes of the Ethics Canvas in the form of risks and barriers to be taken up by the project team and to discuss and commit their mapping in the roadmap (30 min)
- Derivation of related actions to prevent or mitigate the previously identified risks and barriers to directly link them (45 min)
- Discussion and approval of the overall outcome (10 min)

Broadly conceived, the workshop outcomes show that there may be a complex interdependency between technical design decisions and potential stakeholder impacts and concerns. Furthermore, we distinguished between actions that need to be considered prospectively in the case of a market launch and actions that should and, subject to available resources, can be carried out within the project timeframe. Figure 2 presents the entire RRI roadmap as an outlook that needs to be evaluated regarding its impact at both the strategic and operational levels.

In line with the methodological approach, we have only sketched the outcomes of the collaboration. In the next section, we reflect on our procedural approach to implementing RRI in practice and discuss these research results.

## Discussion

After presenting the methodological approach, intermediate and final outcomes, and a first evaluation, we will now discuss our implementation of the RbD framework in the development of an AI-supported exergame in terms of the most dominant themes: (i) its effectiveness and efficiency, (ii) questions of responsibility allocation within interdisciplinary teams, and (iii) the challenges of the translational research context. Where appropriate, we will comment on avenues for further improvement.

### Effectiveness and efficiency of the process

The scope and quality of outcomes from our approach demonstrate that we have implemented a feasible process to create a strategic basis for considering RRI in practice. The iterative and participatory procedure enabled the project team to establish a common basis for communication and ensured that everybody was aware of the ethical points of discussion. The Ethics Canvas method and the 60-min duration of the first workshop were sufficient for the participants to capture a wide range of relevant ethical aspects of exergames and AI, as evaluated in the related wrap-up section. The ethical reflection did not prioritise issues at this stage. However, the project team was able to focus on the issues relevant to the intended product rather than being too general.

In contrast, the formulated RRI vision does not refer to the core RRI tenets but focuses singularly on the well-being of the users. Even when limited to the RRI product dimension, the vision does not acknowledge, e.g., the relevance of physiotherapists or the societal embedding of the technology. However, these aspects were addressed in the Ethics Canvas. Furthermore, despite the highly subjective goal of working towards user well-being, the RRI vision does not mention the RRI process dimension or its principles of inclusivity and responsiveness, from which actions can be derived and legitimised. Again, this is in stark contrast to the breadth of issues and inclusive means of addressing them that the collaboration has derived. A stronger link between the ethical outcomes from the first workshop and the second would have supported the transfer of a more holistic RRI dimension into the vision. This is important because the RRI vision is the tool by which the project team could assess the ethical alignment of actions throughout the development process. We, therefore, suggest at least one additional iteration to further discuss and review the RRI vision after the roadmap has been developed to reflect on and appreciate the guiding function of the vision and its usefulness as a quality management tool.

A limitation of the results is that most of the social implications of the intended product are purely hypothetical and need to be confirmed or refuted by the users and physiotherapists. This barrier is commonly encountered when implementing RRI practices, as ethical considerations are often based on assumptions and possible scenarios, thereby appearing less concrete due to the risks being framed as merely potential or overly exaggerated (McLennan et al., 2022). However, this is where the usefulness of the RbD approach becomes apparent. If challenges, risks and barriers can be assessed as plausible but essentially uncertain, then the appropriate RRI actions to be considered would aim to reduce this uncertainty. Based on our observations and the RRI roadmap produced, the project team has followed this rationale and, therefore recognized the need to implement appropriate RRI actions in the future to confirm or refute hypotheses on ethical issues. We are convinced that increasing the inclusion of external stakeholders directly in the RbD process may help substantiate or refute some perceived ethical implications immediately. In our case, the project team had already invested a significant amount of time that had not been accounted for in the project's grant proposal, so there was an incentive to conduct a first RbD iteration in as lean a format as possible.

Thus, our approach aligns with Kuzma's and Roberts' (2018) suggestion to develop supporting tools and processes in a short-term perspective of RRI to avoid overstraining the project team and its resources. The RbD approach does justice to both: short-term support but clearly advocating the allocation of RRI-related resources later. In fact, our

approach was elaborated in accordance with the requirements of the project. Although a tendency of "pre-formatting" (Zwart et al., 2014) was observed, it was primarily on the process level rather than limitations of ethical analysis. The evaluation indicated that the ethical depth was sufficient, and the proposed adjustments would not make the process unmanageable. Nonetheless, the collaboration began with a delay to the actual project start, causing ethical issues that could have challenged the entire project to be disregarded. The collaboration instead focused on testing a feasible process with supportive methods.

## Responsibilities in interdisciplinary collaboration

Interdisciplinary collaboration arises from the combination of different expertise, but at the same time, it requires considerable coordination. Who participates, what is their role, and who is responsible for what at what time? These are all relevant questions that need to be negotiated.

A key insight we perceived is the benefit of having the entire interdisciplinary project team working together in the workshops. From our experience, we can confirm that the opportunity to involve multiple perspectives, perceptions and levels of expertise in the project broadens the topics touched upon during ethical reflection, clarifies ethical issues at different stages of product development (Gohar et al., 2019; Reijers et al., 2018), and acknowledges the often-diffused responsibility (McLennan et al., 2022). Thus, in our approach, interdisciplinary collaboration is assumed from the outset—at least for the ethical context analysis—in contrast to Rommetveit et al.'s (2019) perspective that sees interdisciplinarity as an outcome to be achieved. At the same time, the involvement of the entire team imposes a significant aggregate demand on the resource of time. At a minimum, one could aim for a group with representatives from all hierarchy levels and project expertise.

Another notable insight from our collaboration relates to the composition of the collaboration. The interaction between the developers, the embedded ethicist and the external technoethicist, allowed the expertise of the product dimensions to be intertwined with the ethical dimension. This is an essential aspect because a separate analysis by an ethicist would likely result in findings that are not as close to the specific product context while having an embedded ethicist strengthens the ethical depth of the analysis and can further help moderate the open, interdisciplinary mindset and basic ethical framework for this process (McLennan et al., 2022). Throughout this collaboration, in addition to providing ethical expertise, the embedded ethicist was also responsible for tailoring the RbD process to the constraints and requirements of the project, as well as preparing and moderating the workshops. Thus, the role of the embedded ethicist was extended, as embedded ethics envisages the integration of ethics throughout the development process and the creation of transparency (McLennan et al., 2022) but does not say anything about being responsible for the appropriate process introduction for ethical analysis. This carried the risk of a trade-off when changing roles, particularly in terms of ethical depth, as ensuring compliance with the process was essential to achieving the purpose of the collaboration—facilitating the project team's own competencies in terms of a cultural shift towards making project teams more responsible and trustworthy. In interdisciplinary research, this possible disadvantage in one's research topic, in this case, that of the embedded ethicist, in comparison to separate accompanying research instead of collaboration, is well known but is countered by interdisciplinary translational researchers who state that their goal is impact (Gohar et al., 2019). However, we managed to maintain ethical depth with the additional support of the external technoethicist at relevant stages of the process. Accordingly, we are convinced that this approach alleviates the challenge identified by Zwart et al. (2014) regarding the fruitful proximity and necessary autonomy of integrated ethics. We achieved balance by including an ethicist within the project team and engaging an external technoethicist for an objective perspective. This allowed for critical discussions without affecting the team's collaboration and mood, which could have been detrimental if left solely to the embedded ethicist.

Nevertheless, it should be noted that the external technoethicist is not included in this project but is part of the broader project ecosystem, which has yet to be assessed regarding any friction between objectives. Furthermore, this constellation of ethicists fulfils McLennan et al.'s (2022) criterion that ethicists possess domain expertise. However, the impact of this feature is expected to be most evident at the operational level, particularly in debates regarding the ethical assessment of technical solutions, as this could pose a challenge for ethicists without some technical understanding.

Preparing an RRI roadmap at the outset of collaboration facilitates the identification of potential ethical issues surrounding the innovation. It enables the assignment of practical tasks to be addressed throughout the development process. This prompts questioning whether this strategic approach can supplant or enhance McLennan et al.'s (2022) "gold standard" of embeddedness, as ethics is already regarded as an integral component of the development process. As such, this proactive planning may reduce the necessity of identifying ethical issues during team meetings. Nonetheless, regular exchange can be beneficial as it aids communication for genuine collaborative interdisciplinarity. It avoids the impression that the ethicist is a mere external challenger of the project or a service provider.

Reflecting on the types of (non-)embeddedness of ethicists, we conclude that our approach prioritises

interdisciplinary collaboration for integrating ethics into innovation projects rather than determining their level of embeddedness. We have temporarily used the term "embedded ethicist" for our collaboration, but not in a broader context of being embedded in the project team. Therefore, this case study is an entry point for embedded ethics as it showcases a strategic plan to integrate ethical considerations throughout development.

In addition to the roles, the question of responsibility needs to be clarified and assigned. We want to emphasise the overall relevance of the discussions and especially the agreement on the intermediate outcomes. In our implementation, we wanted to obtain proper consent from the whole project team to ensure a complete comprehension of the ethical matters incorporated within the RRI roadmap, heightening awareness and commitment towards addressing said issues. However, the lack of time for in-depth discussions often meant that consent was expressed implicitly, i.e., by not contradicting each other, rather than explicitly agreeing and committing. The effects may be similar, but the explicit agreement would probably result in a more substantial commitment and sense of ownership on behalf of the whole project team, the project leader or those ultimately responsible. Therefore, we suggest explicit checkpoints for approval of the intermediate outcomes to increase the awareness and importance of the outcomes for further processing. This also applies to the derived RRI actions to achieve a meaningful transfer of strategic insights directly into operational implementation. In this context, the check for consensus could enhance the scrutiny of ethical issues and facilitate further discussion. It remains unclear to what degree this strategic planning level might create friction, as it is a self-binding instrument for strategic preparation of ethical aspects, it can be deviated from the plan at any time, especially if the implementation of actions is not manageable.

Further case studies may be required to investigate the implementation of RRI on both strategic and operational levels to gain insights into addressing these issues and to assess their impact and effort better. This also raises questions of accountability, which were not particularly examined in this case study. However, adherence to a certifiable standard like this can ensure compliance with an established RRI roadmap. Hence, consensus is an aspect that could have been made more transparent in our process, as responsibilities were only implicitly assumed for RRI actions. Looking ahead, the forthcoming inclusion of user interests will not only raise questions of responsibility. Instead, we suspect that the balance of power in decision-making will be relevant in the case of conflicts, as it is largely unclear how decisions are made in these situations (Goirand et al., 2021). Consequently, similar to McLennan et al. (2022), we suggest assigning responsibilities and establishing a decision-making structure from the beginning.

## Translational context

The context of translational research is characterised by collaborations between industry and academia, typically supported by grants of limited duration. These project constellations are bound to dissolve at some point. Thus, in contrast to the implementation of the RbD process in industry, the RRI roadmap in the translational context, which is a central outcome of the collaboration, runs the risk of referring to a configuration of technical skills, expertise and, ultimately, responsibilities that will no longer exist once the collaboration ends. While such configurations can be usefully extrapolated into the future within a pure industrial framework, in the context of translational research, it may not be possible to assume that academic or industrial partners will be able to muster the resources to sustain the collaboration, or even a new functional one, that can support the RRI vision and RRI actions outlined in the roadmap. Accordingly, responsibilities for actions need to be explicitly assigned at the latest after the completion of the roadmap—as suggested above, ideally at the very beginning of the collaboration. Ultimately, tensions may arise, as the most likely partner to be able to assign responsibility in a meaningful way would be the industrial partners. Therefore, these will consider a potentially necessary build-up of additional expertise within the company itself or appropriate funds to temporarily acquire it (Auer & Jarmai, 2018).

A particular challenge in this project was the asynchronous timing, perhaps typical of grant-based research. While the project started immediately, the collaboration with the platform could not start until about a year into the project due to delays in hiring and negotiations with other projects. Especially in larger consortium projects, where RRI support is performed as a platform aspect within a multi-actor ecosystem (Tsujimoto et al., 2018), the initiation phase for RRI collaborations must be considered. As a result, the roadmap was not created at the beginning of the project. Therefore, participatory ethical reflection and mapping into a project development roadmap requires foresight to enable sustainable effects. However, even though the RRI roadmap was developed in the middle of the development process, and thus constraints on the implementation of actions were recognised, the RbD process still demonstrates the opportunity of mapping potential impacts with a long-term perspective, even after the funding of a project has ended.

Another takeaway relates to the subject matter of our collaboration. While in the translational context, technologies are often at the cutting edge of science, the relevant ethical aspects are not necessarily specific to them. For example, even in AI-related projects, some key risks and challenges are not specific to AI. Of course, this does not devalue any work to mitigate them but rather demonstrates the benefit

and necessity of iterative, interdisciplinary collaboration aimed at anticipating and reflecting on the ethical implications of the project's ecosystem rather than, for example, applying only specific guidelines that might focus on only parts of the product. Such an approach would risk overlooking the broader project context and its requirements. This is why we believe that our approach allows the impact of the effort not to be limited to the trustworthiness of AI components but rather to enable a shift towards a culture of trustworthiness inherent to the entire innovation process. For this reason, we promote the need for contextualisation rather than a one-size-fits-all approach (Goirand et al., 2021; Rommetveit et al., 2019) to ensure the credibility of the effort and highlight the opportunities of the RbD process (as well as associated implementation methods such as the Ethics Canvas) as transferable methods to support interdisciplinary collaborations aimed at translating RRI into practice.

## Conclusion

The project team's objective of collaboration was achieved as we identified and analysed the ethical implications of the project throughout the development process and elaborated actions to address these implications. The RbD process proved to be a practical basis for tailoring interdisciplinary collaboration to consider ethics and implement RRI in a translational research and development process, as ethical depth could be ensured significantly. In particular, the opportunities from involving the entire project team were highlighted. However, questions about the allocation of responsibilities in interdisciplinary collaboration, the roles—especially concerning frictions of the embedded ethicist in this setting—and the challenges of the translational context were discussed, e.g., additional iterations to ensure alignment with an overall vision and to significantly ensure transparent decision making.

We are convinced that our approach provides insights at two levels: first, it shows how RRI can be implemented in a translational context, providing practical information on the methods and workshop designs used to successfully reconcile the requirements of a lean process with ensuring the necessary ethical depth; second, it shows how interdisciplinary collaboration with embedded ethics enables researchers to identify, engage with and plan for ethical issues early in the development process so that RRI can be effectively translated into practice.

With suitable modifications, our approach provides a fruitful case study in facilitating the project team's own competencies in terms of a cultural shift towards making project teams more responsible and trustworthy. Adherence to a certifiable standard is about making the product and innovation process more responsible and making this visible to the outside world—an essential ingredient for project teams to demonstrate their trustworthiness. In conclusion, we believe that our approach serves as a useful case study on which to build in order to further tailor RRI methodologies to the practical requirements of industrial application.

## Appendix

See Table 1.

**Table 1** streamlined version of workshop outcomes with number of voting in brackets

| Identified issues |
|---|
| (1) *Individuals and groups affected:* |
| Identify the individuals, collectives, organisations that can be affected by your product (men/women, user/non-user, professional bodies, competing companies, government agencies…) |
| • Physiotherapists and trainers/coaches (6) |
| • Target group: people between 55 and 67 years (4) |
| • Fitness professionals (3) |
| • Elderly people (3) |
| • People with disabilities (2) |
| • Friends/peers/family members of the users (2) |
| • Competing companies (e.g., sport studios, other home workout companies) |
| • Technical staff (install, maintain, and host the technical infrastructure) |
| • Young, healthy people |
| • Health insurances |
| • Patient associations |
| • Care ward personnel & management |
| • System developers |
| • Self-aid groups |
| • Psychotherapists |
| (2) *Behaviour:* |
| Identify (problematic) changes to individuals' behaviour (e.g., differences in habits, time-schedules, choice of activities, people behaving more individualistic/collectivist, more or less materialistic…) |
| • Even more inhouse time (5) |
| • Unhealthy exercise habits or harm from wrong feedback or instructions (2) |
| • Getting off track due to gamification and not focus on needed exercise (2) |
| • Overreliance on system and devaluation of in-person training (2); self-guided training beyond reasonable stress levels |
| • More time spent on exercising, less time for other things/individualism/hobbies (2) |
| • People are motivated doing sports due to gamification (1) |
| • Stress from gamification elements (success and failure) (1) and frustration (1) |
| • Social distance (e.g., no group sport or no personal interaction) (1) |
| • Hurdles to train in more familiar or relaxing environments (e.g., outside) (1) |
| • Change of caretakers' and physiotherapists' job profile (learn to use and debug system) |
| • Preference of reduced personal contact (physiotherapists and users) |
| • Change in schedule needed to make time for the exercise |
| • Disclosure of personal health information (e.g., to the virtual agent) |
| • More healthy and more frequent exercise possible |
| • More freedom in self-management (e.g., no more dependency of fitness centre or physicians date) |
| • Decrease of motivation due to responsibility of self-management (e.g., challenge for depressive or stroke patients where a prescribed date structure is useful) |
| • Negative effect on body awareness from technologically determined pose and motion evaluation |
| (3) *Relations and groups conflicts:* |
| Identify (problematic) differences in individual behaviour and the impact on the relationship between the groups identified |
| • Conflict between therapist and user (e.g., different intention of use) (3) |
| • Social isolation (e.g., doing sports alone) (3) |
| • Social side effect of therapy may be diminished (3) |
| • Less personal interaction between patients and therapists (2) |
| • Reduced affordability due to expensive technical infrastructure/hardware (2); potentially no reimbursement by health insurances |
| • Impact on body awareness (people rely on feedback on screen only) (2) |
| • Competitive design might lead to conflicts (e.g., over-exhaustion, competitive behaviour) (1) |
| • Users might want to get reimbursed for using the system, but health insurance might not want that (1) |
| • Mistrust of technology and instructions (1) |
| • Inability to customize one's own training schedules to best meet the need (1) |
| • Need for private coordination if used in a household (no interruptions) (1) |
| • Lack of direct physical support in movement |
| • Conflicts due to need for explanation if not fully self-explaining in usage |
| • Possible disadvantage for people not wanting to use the product (e.g., less access to therapy) |

**Table 1** (continued)

| Identified issues |
|---|
| (4) *Worldviews:*<br>How can the general perception of a person's role in society be affected by the project?<br>• Effect on role of physiotherapists (need for social interaction vs. non-human automatic feedback) (5)<br>• Shift in authority and requirement of new forms of trust due to premise of technology can better judge the "correctness" of human motion (5)<br>• Establishment of "fashionable" or "trendy" exercising also among older age groups (4)<br>• Elderly people as a new target group for sports sector (2)<br>• Change in role of the physiotherapists (e.g., less treatment of everyday life suffers; rather in the role of explaining the product than in personal training) (1)<br>• Perception as innovative clinic or hospital providing the system (1)<br>• Negative perception from elderly people about other people using a robot (1)<br>• Staying indoors more instead of going out to do exercise<br>• Decrease of the need for professional physiotherapy on everyday life suffer, (e.g., caused by a desk job) |
| (5) *Product or service failure:*<br>Identify potential negative impact of your product or service failing to operate as intended (technical or human error, security breach, data loss…)<br>• Incorrect detection and motion correction causing harm to the users (8)<br>• Diversity and algorithmic bias (e.g., different appearance of users from training data; bias in dataset) (5)<br>• Injuries from wrong feedback or performance (2)<br>• Data breach (e.g., 3D data as identifier for the person) (2); data loss of large amount of 3D data<br>• Decrease in motivation by data loss (e.g., loss of tracked progress) (2)<br>• Effect of a strong line between correct and incorrect on functionality (e.g., step by step mobilisation) (1)<br>• Potential fire hazard due to battery of robot<br>• Co-morbidity in terms of psychological and physiological conditions (harmful exercises)<br>• Dissatisfaction of health outcomes if agent appears more professional than it actually is |
| (6) *Problematic use of resources:*<br>Identify possible negative impacts of the consumption of resources of your project (climate impacts, privacy impacts, employment impacts…)<br>• Privacy issues (local processing of data, transparent use of external data processing/cloud storage) (6)<br>• Privacy impact of visible exercise data (3)<br>• Requirement of human resources to host, run and maintain the system (3)<br>• Physiotherapists may feel threatened at the prospect of AI based training support, but the ability of the system to augment their practices might only need to be tailored to better suit their and their patients' needs (3)<br>• Hardware / energy needed (2)<br>• Fully tech based, everything manufactured and electric usage (1); conflict minerals<br>• No easy reparation (1)<br>• Potential of malicious options for misuses of data (e.g., psychometric analyses) (1)<br>• By-products from motion analysis (e.g., detection of injuries or physical problems)<br>• Relief of physiotherapists and health system (e.g., less need for human resources) |

**Acknowledgements** The authors would like to express their gratitude to the workshop participants of the project team for their valuable and constructive contributions to the procedure. Special thanks to SZENARIS partners Prof. Dr. T. Stadie and M. Lampe for their close cooperation.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) through the KI-SIGS - Project (FKZ: 01MK20012B).



## References

Auer, A., & Jarmai, K. (2018). Implementing responsible research and innovation practices in SMEs: Insights into drivers and barriers from the Austrian Medical Device Sector. *Sustainability, 10*(1), Art. 1. https://doi.org/10.3390/su10010017

Benzing, V., & Schmidt, M. (2018). Exergaming for children and adolescents: Strengths, weaknesses, opportunities and threats. *Journal of Clinical Medicine, 7*(11), 422. https://doi.org/10.3390/jcm7110422

Bhattacharya, I., Patil, V. S., & De Sousa, A. (2022). Ethical issues in the use of exergames in the elderly. *Global Bioethics Enquiry Journal, 10*(2), 75–79. https://doi.org/10.38020/GBE.10.2.2022.75-79

Blok, V., Iñigo, E. A., Novelskaitė, A., Pučėtaitė, R., Adams, A. A., Wakunuma, K., Jiya, T., Grabner, L., Lorenz, L., Sánchez de Madariaga, I., & Novella, I. (2020). Recommendations for the development of a competitive advantage based on RRI. Zenodo. https://doi.org/10.5281/zenodo.4972691

Breyer, S., & Herzog, C. (2022). Integrating ethical considerations into innovation design. In M. Friebe (Ed.), *Novel innovation design for the future of health: Entrepreneurial concepts for patient empowerment and health democratization* (pp. 253–282). Springer International Publishing. https://doi.org/10.1007/978-3-031-08191-0_23

Burget, M., Bardone, E., & Pedaste, M. (2017). Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and Engineering Ethics, 23*(1), 1–19. https://doi.org/10.1007/s11948-016-9782-1

Business Model Foundry AG. (2014). Business Model Canvas. Business Model Foundry AG.

CEN CWA 17796:2021. (2021). *Responsibility-by-design—Guidelines to develop long-term strategies (roadmaps) to innovate responsibly* (Standard CEN CWA 17796:2021). https://standards.iteh.ai/catalog/standards/cen/8e3cfe68-8449-49f8-b87c-d3efe20da158/cwa-17796-2021

Demers-Payette, O., Lehoux, P., & Daudelin, G. (2016). Responsible research and innovation: A productive model for the future of medical innovation. *Journal of Responsible Innovation, 3*(3), 188–208. https://doi.org/10.1080/23299460.2016.1256659

Dreyer, M., Chefneux, L., Goldberg, A., von Heimburg, J., Patrignani, N., Schofield, M., & Shilling, C. (2017). Responsible innovation: A complementary view from industry with proposals for bridging different perspectives. *Sustainability, 9*(10), 1719. https://doi.org/10.3390/su9101719

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4people-an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*, 689–707. https://doi.org/10.1007/s11023-018-9482-5

Floridi, L., & Strait, A. (2020). Ethical foresight analysis: What it is and why it is needed? *Minds and Machines, 30*(1), 77–97. https://doi.org/10.1007/s11023-020-09521-y

Gohar, F., Maschmeyer, P., Mfarrej, B., Lemaire, M., Wedderburn, L. R., Roncarolo, M. G., & van Royen-Kerkhof, A. (2019). Driving medical innovation through interdisciplinarity: Unique opportunities and challenges. *Frontiers in Medicine*. https://doi.org/10.3389/fmed.2019.00035

Goirand, M., Austin, E., & Clay-Williams, R. (2021). Implementing ethics in healthcare AI-based applications: A scoping review. *Science and Engineering Ethics, 27*(5), 61. https://doi.org/10.1007/s11948-021-00336-3

Grunwald, A. (2011). Responsible innovation: Bringing together technology assessment, applied ethics, and STS research. *Enterprise and Work Innovation Studies, 7*, 9–31.

Gurzawska, A., Mäkinen, M., & Brey, P. (2017). Implementation of responsible research and innovation (RRI) practices in industry: Providing the right incentives. *Sustainability, 9*(10), 1759. https://doi.org/10.3390/su9101759

Hallensleben, S., Hustedt, C., Fetic, L., Fleischer, T., Grünke, P., Hagendorff, T., Hauer, M., Hauschke, A., Heesen, J., Herrmann, M., Hillerbrand, R., Hubig, C., Kaminski, A., Krafft, T., Loh, W., Otto, P., & Puntschuh, M. (2020). From Principles to Practice—An interdisciplinary framework to operationalise AI ethics. Bertelsmann Stiftung; VDE.

Herzog, C. (2021). On the risk of confusing interpretability with explicability. *Nature Machine Intelligence*. https://doi.org/10.1007/s43681-021-00121-9

Koene, A. (2017). Algorithmic bias. *IEEE Technology and Society Magazine, 36*(2), 31–32.

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

Li, J., Erdt, M., Chen, L., Cao, Y., Lee, S.-Q., & Theng, Y.-L. (2018). The social effects of exergames on older adults: systematic review and metric analysis. *Journal of Medical Internet Research, 20*(6), e10486. https://doi.org/10.2196/10486

Martinuzzi, A., Blok, V., Brem, A., Stahl, B., & Schönherr, N. (2018). Responsible Research and Innovation in Industry—Challenges, insights and perspectives. *Sustainability (switzerland)*. https://doi.org/10.3390/su10030702

McLennan, S., Fiske, A., Tigard, D., Müller, R., Haddadin, S., & Buyx, A. (2022). Embedded ethics: A proposal for integrating ethics into the development of medical AI. *BMC Medical Ethics, 23*(1), 6. https://doi.org/10.1186/s12910-022-00746-3

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available ai ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 26*(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Morrison, M., Mourby, M., Gowans, H., Coy, S., & Kaye, J. (2020). Governance of research consortia: Challenges of implementing Responsible Research and Innovation within Europe. *Life Sciences, Society and Policy, 16*(1), 13. https://doi.org/10.1186/s40504-020-00109-z

Nazarko, L. (2020). Responsible research and innovation in enterprises: Benefits, barriers and the problem of assessment. *Journal of Open Innovation: Technology, Market, and Complexity, 6*(1), 12. https://doi.org/10.3390/joitmc6010012

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

Pacifico Silva, H., Lehoux, P., Miller, F. A., & Denis, J.-L. (2018). Introducing responsible innovation in health: A policy-oriented framework. *Health Research Policy and Systems, 16*(1), 90. https://doi.org/10.1186/s12961-018-0362-5

Petersen, E., Potdevin, Y., Mohammadi, E., Zidowitz, S., Breyer, S., Nowotka, D., Henn, S., Pechmann, L., Leucker, M., Rostalski, P., & Herzog, C. (2022). Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions. *IEEE Access, 10*, 58375–58418. https://doi.org/10.1109/ACCESS.2022.3178382

Porcari, A., Pimponi, D., Borsella, E., & Mantovani, E. (2019). PRISMA RRI-CSR Roadmap.

Reijers, W., Koidl, K., Lewis, D., Pandit, H. J., & Gordijn, B. (2018). Discussing ethical impacts in research and innovation: The Ethics Canvas. In D. Krebs, C. Ess, L. Leenen, & K. Kimppa (Eds.), *This changes everything—ICT and climate change: what can we do?* IFIP Advances in Information and Communication Technology. https://doi.org/10.1007/978-3-319-99605-9_23

Ribeiro, B., Bengtsson, L., Benneworth, P., Bührer, S., Castro-Martínez, E., Hansen, M., Jarmai, K., Lindner, R., Olmos-Peñuela, J., Ott, C., & Shapira, P. (2018). Introducing the dilemma of societal alignment for inclusive and responsible research and innovation. *Journal of Responsible Innovation, 5*(3), 316–331. https://doi.org/10.1080/23299460.2018.1495033

Rommetveit, K., Dijk, N., Gunnarsdóttir, K., O'Riordan, K., Gutwirth, S., Strand, R., & Wynne, B. (2019). Working responsibly across boundaries? Some practical and theoretical lessons. *International*

*handbook on responsible innovation* (pp. 83–100). Edward Elgar Publishing. https://doi.org/10.4337/9781784718862.00012Rüth

Rüth, M., & Kaspar, K. (2021). Educational and Social Exergaming: A perspective on physical, social, and educational benefits and pitfalls of exergaming at home during the COVID-19 pandemic and afterwards. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.644036

Stahl, B. C., Akintoye, S., Bitsch, L., Bringedal, B., Eke, D., Farisco, M., Grasenick, K., Guerrero, M., Knight, W., Leach, T., Nyholm, S., Ogoh, G., Rosemann, A., Salles, A., Trattnig, J., & Ulnicane, I. (2021). From Responsible Research and Innovation to responsibility by design. *Journal of Responsible Innovation, 8*(2), 175–198. https://doi.org/10.1080/23299460.2021.1955613

Tobaigy, A., Alshehri, M. A., Timmons, S., & Helal, O. F. (2018). The feasibility of using exergames as a rehabilitation tool: The attitudes, awareness, opinions and experiences of physiotherapists, and older people towards exergames. *Journal of Physical Therapy Science, 30*(4), 555–562. https://doi.org/10.1589/jpts.30.555

Tsujimoto, M., Kajikawa, Y., Tomita, J., & Matsumoto, Y. (2018). A review of the ecosystem concept—Towards coherent ecosystem design. *Technological Forecasting and Social Change, 136*, 49–58. https://doi.org/10.1016/j.techfore.2017.06.032

van de Poel, I., Asveld, L., Flipse, S., Klaassen, P., Scholten, V., & Yaghmaei, E. (2017). Company strategies for responsible research and innovation (RRI): A conceptual model. *Sustainability, 9*(11), 2045. https://doi.org/10.3390/su9112045

von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 51–74). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118551424.ch3

Wachter, S., Mittelstadt, B., & Russell, C. (2020). Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3547922

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. *AIES*. https://doi.org/10.1145/33066183314289

Yaghmaei, E., & van de Poel, I. (Eds.). (2020). *Assessment of responsible innovation*. Routledge. https://doi.org/10.4324/9780429298998

Zwart, H., Landeweerd, L., & van Rooij, A. (2014). Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI.' *Life Sciences, Society and Policy, 10*(1), 11. https://doi.org/10.1186/s40504-014-0011-x