**Title:** The future of reproducibility for researchers, funders and publishers

**Description:**

Futures studies that explore the future of reproducibility from different stakeholders and perspectives with the aim to identify the steps needed to make reproducibility across contexts more tangible and probable and provide pathways and give recommendations to facilitate this future.

**Contributors:**

Joeri Tijdink, Serge Horbach, Tony Ross-Hellauer, Nicki Lisa Cole

**Affiliated institutions:**

AmsterdamUMC, location VUmc, KNOW Center, Aarhus University

**Subjects:**

4 stakeholder groups: researchers/experts in reproducibility from computer science (machine learning), researchers/experts in qualitative research (social sciences), publishers and funders.

**Tags/keywords:**

Reproducibility, futures studies, replication, replicability, epistemic diversity, Horizon Europe

1. Abstract

2. Preamble/rationale

Futures studies is the systematic study of possible, probable and preferable futures. It aims to explore and predict the future by envisioning alternative future scenarios, which then inform strategies for action to eventually shape the desired future at different levels. This can be done for individuals, companies, systems or even at global and societal levels (1-3). In the words of Inayatullah (2013), ‘Futures studies create alternative futures by making basic assumptions problematic. Through questioning the future, emerging-issues analysis and scenarios, the intention is to move out of the present and create the possibility of new futures.’

**Why do we use futures studies in the TIER2 project?**

[TIER2](https://tier2-project.eu/) is an EC Horizon Europe-funded project (2023-2025) that aims to boost knowledge on reproducibility, create tools, engage communities, implement interventions and policy across different contexts to increase the reproducibility, re-use and overall quality of research results (4). Key to the project is to explore how these interventions are best fitted to the diverse epistemic, cultural and sociotechnical specificities of different domains of research (5).

The exploratory and purposive methods commonly used in futures studies are an excellent tool by which to begin this investigation. Futures studies can challenge current frameworks and commonplace assumptions by creating scenarios of the future (both desired and not) that can inform fundamental rethinking of the necessary steps required for their realisation (in the case of desired futures) or avoidance (where not desired).

Furthermore, futures studies can be used to test and reflect on existing policies and explore whether other policies should also be considered and how the present state of reproducibility integration in academia can adapt with or evolve towards potential alternative futures. Creating and testing potential scenarios of how reproducibility will evolve helps to prepare and respond to the challenges that will be faced in these futures. Part of the TIER2 project is to add to existing policy that aims to improve reproducibility with new insights in how to promote reproducibility and what are potential barriers in the future.

In short, using futures studies for policymaking, planning, creating and testing tools/practices and exploring potential barriers and facilitators for fostering reproducibility will develop insight into the future of reproducibility. In our approach, a series of scenario workshops with four different stakeholder groups will be set up to we should forecast the future of reproducibility, then we should compare these scenarios and ask ourselves whether other out-of-the –box or alternative scenarios may be warranted for the future.

**Methods of Futures Studies: the six pillars:**

In this futures studies project, we aim to determine and shape the future of reproducibility. For this, we think that the 6 pillars of futures studies expert Inayatullah (2) can help us to design the exercises and use this as a conceptual framework. His work created these six pillars building on Dator’s Manoa School (2, 6). We foresee that these pillars can guide the foresight process. The framework provides several methods and tools that are applicable to the aim of creating the future of reproducibility.

The 6 pillars are a general theoretical framework for futures thinking that is linked to methods and tools. These pillars are *1) mapping, 2) anticipation, 3) timing, 4) deepening, 5) creating alternatives, 6) transforming*. As proposed by Inayatullah (2) we adapt this general framework to the conditions and objectives of the TIER2 project, resulting in the approach sketched out below. In our study, we will use 3 pillars. First, we *map* the current state of reproducibility (1st pillar), then deepen the future by *creating alternatives* (5th pillar) and start exploring the path to *transformation* (6th pillar) by deepening our ideas about facilitators and barriers on the road to transformation.

Described by Boulding (7), back casting works by moving individuals into the preferred future and asking them what happened in the last ten years to bring them to today. Doing this makes the preferred future more concrete and tangible. This technique also focuses on the potential facilitators and barriers that will lead the road in the upcoming ten years.

The reason for selecting these 3 pillars are multiple. First, we think that these are the most important pillars that will help us reach our goals and aims of our study. Mapping is necessary to understand the present state of reproducibility and get participants to think about reproducibility in general. Secondly, we wanted to cocreate different scenarios with the participants. These scenarios are helpful in understanding what direction reproducibility should be taking and how this can be accomplished. The final pillar, transformation, is often considered the cornerstone of futures studies. Since we think that the transformation process is essential to understand the future and to tackle potential barriers, it is pivotal to back cast in order to understand the road towards the preferred future. Furthermore, our focus on these three pillars by no means implies that the other pillars are not implied in our work. Instead, we design three sets of exercises for our workshops, tailored to these three pillars, that integrate elements of other pillars (e.g. deepening of the scenarios) and only omit aspects of pillars that are of less concern to our overall project (mainly pillars 2 and 3).

### **2. Study Information**

**Research question:** What will/should be the future of reproducibility in four domains, according to experts and stakeholders in these specific fields (machine learning (part of computer sciences), qualitative research (social sciences), funders, publishers). This is specified in several more specific questions that we use to generate the data:

Sub 1: What trends and events can be identified that have created the present state of reproducibility? (pre-workshop survey)

Sub 2: What is the predicted future of reproducibility? (pre-workshop survey)

Sub 3: What are alternative futures, feared ones? Dystopian ones?

Sub 4: What is your preferred future of reproducibility?

Sub 5: What are the steps to come to this future based on the present?

Sub 6: What facilitators and barriers will come on our way towards the preferred future? (validation workshop)

**Aim:** the aim is to determine the future of reproducibility. We do this by exploring potential scenarios of the future, examine what steps are needed to make these scenarios workable in the future and detect barriers and facilitators for reproducibility in the current research system and the future system. In addition, we also aim to identify the key factors and the key stakeholders that may have the ‘power’ to influence these scenarios.

**Hypotheses:** no hypotheses are formulated due to the explorative nature of the study

**Anticipated Duration**: Nine months. See our timeline below for details.

### **3. Design Plan**

**Study type:** Futures studies with four (scenario) workshops. The future of reproducibility will be explored with researchers and experts in the field of reproducibility in machine learning (computer sciences), researchers and experts of qualitative research with some understanding of reproducibility issues (predominantly from the social sciences), and two research stakeholders (research funders and publishers). We aim to look at the future of reproducibility in ten years. We estimate that ten years will be close enough to make the scenarios as concrete as possible and will give room to think about the steps that need to be taken to realise this future.

**Study Design**: Scenario workshops, with a pre-workshop survey to help participants develop their ideas prior to the workshop, and a validation workshop to collectively reflect on and validate initial findings.

**Pre-workshop survey:**

In this survey, we let the participants familiarise themselves with the topic of reproducibility, the goals of the workshop, the exercises, and give them a glimpse of the MIRO board that will serve as the collaborative platform for the workshops. We will also collect demographic data (gender, academic rank, disciplinary field, main research methods used, age, expertise, institution, country) and provide the participants with an informed consent form and information about our project. This format furthermore allows participants to start thinking about the present and potential future state of reproducibility.

The four main survey questions asked are:

1. What is the future of reproducibility if the present scenario continues?

2 What is, according to you, the preferred future of reproducibility in your field/disciplinary domain?

3. What are potential barriers that you expect might make it difficult to reach this preferred future?

4. What could be potential factors to facilitate this preferred future?

**Workshops:**

The six pillars described by Inayatullah (2) will be deployed throughout the four workshops (one for each stakeholder group), which are structured as follows.

We plan for online workshops to last 3,5 hours, with various breaks in between. The four workshops will be organised using the platform MIRO and the communication software ZOOM. Each workshop starts with a round of introduction, so participants get to know each other. Subsequently, the workshops will be organised based on three substantive sessions, each containing several exercises (see below for details). Many of the exercises will have a diverge/converge structure, where potential scenarios are briefly explored first, and are clustered or selected next. This can include, for example, an exercise to think about as many ideas as possible, reflect on the scenarios as a group, and then make a choice as to which scenarios are the best ones to develop, predict and backcast with (see above some more info on backcasting). The role of the facilitator is to guide the participants throughout the process and let them reflect on what scenarios they came up with and on what the group decided to continue with.

**Facilitation**

Each workshop will be guided by an experienced facilitator (JT), and supported by a co-facilitator. The role of the facilitator is to guide the workshops, explain all steps and exercises, and ask follow-up questions on answers given. The role of the co-facilitator is to handle any technical issues, make notes and observations and aid the facilitator.

1. *Introduction exercise*

Introduction of participants, introduction of topic and goals, presentation of the results of the pre-workshop survey, engage the participants and familiarize the participants with the methods of futures studies workshops. Ground rules are: 1) Ask each other open questions; 2) Listen to each other; 3) Feel free to ask the facilitators anything when you're in doubt. We're all learning; 4) All ideas are interesting and welcome. The goal of the workshop is to come up with new ideas.

Outcome: participants get to know each other and the facilitators and learn how to use the MIRO board

Using the MIRO board: Yes, with several short exercises about how to use the post its, how to move around and how to get an overview of the exercises.

Time: +/- 20 minutes

Role of the facilitator: leading the exercises

1. *Mapping exercise (pillar 1 of Inayatullah).*

Participants will give their first take on reproducibility in their field by responding to this question: What is the state of reproducibility in my field/area/discipline? To determine one’s future, one must look at the past and present. We ask participants to think about their most recently finished project and state:

(i)             what reproducibility practices they engaged in,

(ii)            what services/infrastructures they used,

(iii)           how they feel their practices are perceived/rewarded/supported by other stakeholders (researchers, publishers, funders)

After having them individually write these things on a post-it, participants are given some time to reflect on each other’s notes and to add to their own, in case they got inspiration from the other’s.

Subsequently, they will discuss their input, which allows them and the facilitator to elaborate or ask for clarification. Also known as the futures triangle (past – present – future), this is necessary to move towards the future, only by knowing the past and present. With the exploration of the past and present, we ask the participants to write what they would have liked to do in an ideal world, so, in resemblance to the above, (i) what they would ideally have done, (ii) what infrastructures they would ideally have used, (iii) how their practices would ideally have been perceived.

Respondents again write down their thoughts on post-its, share ideas with the group, and discuss the outcomes of the others. The latter includes discussing why these ‘ideal’ practices/infrastructures/perceptions are not the actual ones. This forms a bridge to the next round (where we draft the scenarios).

*Outcome:* overview of the past and present of reproducibility according to the 4 stakeholder groups.

Using the MIRO board: this will be done on the MIRO board

Time: +/- 60 min

Role of the facilitator: energising participants to actively contribute to the MIRO board and writing post-it notes; make sure that each participant has the opportunity to contribute.

1. *Scenario planning (Pillar 5, creating scenarios).* The description of how scenarios may look like (we call this scenario planning in this study) is the core of most futures studies. They open up the present, contour the range of uncertainties, reduce risk, offer alternatives, create more flexible institutions that are ready for change and predict the future. In our approach, participants have to create different scenarios per domain/stakeholder group. The first scenario is *the preferred*. This describes the reproducibility world we want. The second scenario is *the disowned*. This is the world we reject or do not want to negotiate. The third is the integrated scenario where both the preferred and disowned are combined and gives a perspective that may be the most likely scenario. Participants will be stimulated to give as much details as possible on these scenarios. As a final scenario, participants also outline *the outlier* scenario. This is a separate scenario that describes out-of-the-box ideas. This will not be elaborated on when the scenarios are highlighted in the discussions.

*Outcome: four* different scenarios (the preferred, the disowned, the likely, and the outlier scenario) from the participants will be described and further explored on the MIRO board.

Using the MIRO board: the exercises will be done on the MIRO board

Time: +/- 60 min

Role of the facilitator: energising participants to actively contribute to the MIRO board and writing post-it notes; make sure that each participant has the opportunity to contribute.

1. *Backcasting the preferred scenario(pillar 6, transformation).* Working with the participants’ preferred and most likely future of reproducibility, we will ask them to identify what will have needed to happen over the ten years prior to these future scenarios in order for them to be realised. Participants will first work in pairs and then present to each other their outcomes plenary. During this exercise we will help the participants to identify and resolve potential conflicts that might occur over these ten years, and to identify potential barriers and facilitators along the way.

*Outcome:* an overview of the steps that are necessary to come to the preferred future and an overview of the facilitators and barriers that are on the path towards this future.

Using the MIRO board: the exercises will be explained and conducted on the MIRO board

Time: +/- 60 min

Role of the facilitator: energizing participants to actively contribute to the MIRO board and writing post it notes; make sure that each participant has the opportunity to contribute.

At the end of each workshop, the participants are asked to reflect on the aim of the workshop through a last exercise. This last exercise will ask participants to evaluate the end products and let them write down on the MIRO board what they think of the workshops, what went well and how they would evaluate it. After finishing the formal workshop part, we will leave time for the participants and facilitators to continue talking about the topic informally as relevant new insights might still arise (8, 9). Our experience is, however, that in online settings, this will most often result in people leaving soon and reflections on the topic are scarce.

**Validation workshop:**

In this workshop, we invite participants from the four stakeholder groups to reflect on the preliminary findings from their workshop. Every group will receive a short presentation on the results. We will summarise the three scenarios that have been created and present the backcasting steps of the preferred-likely scenario. We also ask them whether we have correctly understood the results and ask them to reflect on implications for the future of reproducibility. We will also check whether the participants have additional points that they want to add to the results. With this information, we can validate our findings, reflect on the results and address potential missing points and ask for ideas that have not been discussed in the workshops. We will answer the following questions in this workshop:

1. Did this summary cover the most important aspects discussed in the workshops?
2. What do you think of this preferred future? (we will ask participants to reflect on this scenario, what are the advantages, the disadvantages, are things missing and what elements of this preferred scenario will be realistic)
3. What are potential barriers and facilitators of this future?

**Pilot test**

Before the first scenario workshop, we will organise a pilot workshop with several colleagues (consortium members from TIER2, OSIRIS, iRISE) to test the different exercises. The participants here are experts in reproducibility. This allows us to examine intended plans and implement adjustments where needed, e.g., for reasons of time. Based on the pilot tests we will make necessary adjustments to the scenario workshops, the questions that we use in the workshop and the study protocol. Material from the pilot tests constitutes part of our empirical data.

**Ethical considerations**

Ethical approval will be obtained from the Ethical Review Board of the Faculty of Humanities at the Vrije Universiteit Amsterdam. Prior to participation, participants are provided with an information letter and privacy policy statement. Participants are asked to sign an online informed consent form.

**Sampling and case selection strategy**

**Participant recruitment:**

For the workshop we will recruit participants based on the following inclusion criteria: people should be working in the specific stakeholder group targeted for the specific workshop (i.e., working in machine learning, qualitative social science, working as a funder, working at a publisher) and have demonstrable expertise and experience with issues of reproducibility in that specific stakeholder group. The latter refers to either studying reproducibility-related issues or contributing to projects or procedures that aim to improve reproducibility or to tackle related issues.

Participants will be recruited through a three-fold recruitment strategy: (i) through existing connections and networks of the TIER2 consortium, (ii) through snowballing, i.e. by asking initially identified potential participants to identify others within their targeted stakeholder group that match the inclusion criteria, and (iii) through internet and literature searches to identify authors of papers, presenters of talks, or contributors to projects related to reproducibility issues within the four target communities. The latter builds on previous work within the TIER2 project.

**Sample size:**

We target to include 6 (range 4-8) participants per workshop.

### **4. Data Collection**

**Data source(s) and data type(s) (required)**

Original qualitative data will be collected. First, we will use the input from the pilot workshop to further develop the exercises and have more knowledge on what we can expect from the participants when asked to describe potential scenarios that shape the future of reproducibility. Participants will provide the data by doing the workshop exercises and completing the pre-workshop survey.

**Data collection methods (required)**

The data are gathered by a) one survey and from different elements of the online workshops; b) the content created on the MIRO board; c) the transcripts that are recorded in ZOOM and written out by Amber script, and d) the observations/notes made by the facilitator and co-facilitator. Amber scripts are created (a secure automatic transcription software, using a European server) and pseudonymized. After audio transcriptions and analyses, audio and video recording will be deleted.

**Data collection tools, instruments or plans (required)**

The survey data are collected with Qualtrics for the pre-workshop surveys. The content created on the MIRO board and the analysis of the transcripts will result in the other materials. Tools for data collection are Qualtrics, the MIRO board and a qualitative analysis software program (e.g. NVivo, MAXQDA). Instruments are the surveys, workshop instructions and validation workshop instructions.

**Stopping criteria (required)**

All workshops, including pre-workshop surveys and the validation workshop, will be carried out and data collection will subsequently be terminated. Given the exploratory nature of the research question and the conduct of only a single workshop per stakeholder group, we don’t think that we will reach data saturation in the four workshops. We expect that new information will be obtained in every workshop and will enrich the outcomes and perspectives of the participants on the future of reproducibility. However, some similarities between the workshops will be explored and compared in the analysis.

**Analyses (see also Lechner, Tijdink et al, 2022 (https://osf.io/tycax)**

When all the workshops are conducted, we will analyse the data, making use of the analysis method which Sanders and Stappers (9) have called ‘analysis on the wall’. In this type of analysis all data are first organised, and subsequently analysed on ‘a wall’; in our case a MIRO board. The steps in this analysis process are explained below.

1. First, we analyse the data from the survey to better understand what our participants actually envision when we talk about the future of reproduciblity.
2. The second step is to organise all data. The data includes all materials made by participants in the workshops (i.e., written content on post-it notes) and the output of the audio transcripts. The audio transcripts will be read by two researchers, independently, and they will mark quotes that stand out and label these with the most important message of the quote. This analysis will be done with qualitative research software (such as MAXQDA or Nvivo). This is a type of inductive analysis used to analyse qualitative research. All identifiable information will be pseudonymized, and all selected quotes of participants will be labelled by workshop name, date and time-stamp. All outputs from the workshops (MIRO board and the main quotes from the transcripts) will be put on a MIRO board.
3. Third, the data will be analysed through an ‘open discussion session’ with the research team. In this session the data from the workshops and the quotes will be analysed through clustering, talking, and reflecting on the findings.
4. Fourth, the findings from the third step will be used to identify themes to map the connections between all items (done per exercise). This will also include thematic analysis on QDA software.
5. Fifth, the validation workshop will use the results of the first 4 steps. This workshop will be used to see whether we have understood the results of the workshops, and check with the participants if we have conveyed their perspectives (a so called membercheck).

**Credibility strategies (required):**

We aim to assure methodological rigour and integrity by writing this preregistration and highlight a couple of strategies that can mitigate potential biases and methodological shortcomings and improve credibility of the findings. These strategies include, member checking of the results of the workshops, including different perspectives from different stakeholders, reflexitivity in both participants and research team, ethical review and the use of the COREQ reporting guidelines.

**Timeline:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Activity** | **Action** | **Time** | **Who?** |
| Stakeholder mapping | * Overview of potential participants for the workshops per stakeholder group (4x) | Jan-March 2023 | All |
| Selecting and inviting participants | * Create list of potential participants * Send out invitations for workshop in June/July | April 2023 | WP4 (JT, SH, TRH, JR) |
| Preparation of workshops | * Send out the surveys to all participants (+ reminders) * Prepare MIRO board with exercises * Conduct pilot test in May to test exercises | April/May 2023 | WP4 (JT, SH, TRH, JR, BL) |
| Conduct of workshops | * Conduct 4 workshops with selected participants | June/July 2023 | WP4 (JT, SH, TRH, JR, BL) |
| Analysis of results | * Use analysis on the wall and deductive content analysis to thematize results * Create scenarios * Send out summary of results to get feedback on results from participants. | July-August 2023 | WP4 (JT, SH, TRH, JR, BL) |
| Validation workshop | * Organise a validation workshop to discuss the results and get feedback | Sept-October | WP4 |
| Finalizing report on futures studies | * Write report for EU * Internal review of the report in month ? * Submit in the portal (M11) | Aug-Nov 2023 | WP4 (JT, SH, TRH, JR, BL) |
| Publishing results | * Write 1 /2 publications on the results * Submit to journal | Sept 2023 – March 2024 | WP4 (JT, SH, TRH, JR, BL) |

### **References**

1. Masini E. Rethinking futures studies. Futures. 2006;38(10):1158-68.

2. Inayatullah S. Futures Studies. Theories and Methods. There's a Future Visions for a Better World. 2013:36-66.

3. Inayatullah S. Tips and pitfalls of the futures studies trade. Foresight. 2000;2(4):369-74.

4. Ross-Hellauer T, Klebel T, Bannach-Brown A, Horbach SPJM, Jabeen H, Manola N, et al. TIER2: enhancing Trust, Integrity and Efficiency in Research through next-level Reproducibility. Research Ideas and Outcomes. 2022;8:e98457.

5. Leonelli S. Rethinking Reproducibility as a Criterion for Research Quality. Including a Symposium on Mary Morgan: Curiosity, Imagination, and Surprise. Research in the History of Economic Thought and Methodology. 36B: Emerald Publishing Limited; 2018. p. 129-46.

6. Press A, editor. The features of Cultures and the Cultures of the Future. New York: Academic Press; 1979.

7. Boulding E, Boulding KE. The Future: Images and Processes: SAGE Publications; 1995.

8. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign. 2008;4(1):5-18.

9. Sanders EBN, Stappers PJ. Convivial Toolbox: Generative Research for the Front End of Design: BIS; 2012.