6. Embrace 'human-in-the-loop:' Human oversight of decision making is a legal requirement for some funders, and a wise choice for all. Human experts and AI systems have complementary strengths, which human-in-the-loop models can help to leverage. The essential criterion for any effective use of AI by funders is that it is trusted by staff, leadership, and applicants, and human-in-the-loop approaches are essential to achieve this.

7. Dedicated oversight structures: While many aspects of AI can and should be covered by existing policies and practices, funders will benefit from having a dedicated structure for oversight of AI applications throughout the organisation. This may come in different forms, such as an internal oversight board or an organisation-wide AI working group. Funders must cultivate appropriate internal expertise to provide oversight and assurance for all AI applications they use.

6.2.3 Principles

8. Responsible AI for funders: Responsible AI frameworks like those discussed in Part 1 are valuable resources, but funders must build on these to develop internal understanding of what responsible AI looks like for each organisation. Funders should work with each other and with internal stakeholders to develop shared understanding of AI accountability, transparency, explainability, agency and oversight, and legitimacy of AI use, as well as how AI affects key concerns for research systems, such as academic freedom and research integrity.

9. Build for the tools of today, not the myths of tomorrow: AI is often presented with far-flung promises of future transformations that are rarely matched by reality. Rather than responding to visions of potential AI futures, funders should focus on tools and technologies currently at hand. These are the technologies that will actively shape research systems and will be most relevant for funders and the researchers they work with. Focusing on current technologies also allows funders to build responsible policies and practices from experimentation and concrete experience with specific use cases.

10. Take a problem-based approach to AI: Funders should approach AI/ML as a toolbox with which to respond to specific problems, rather than seeking out opportunities to make use of AI/ML technologies. A problem-led approach will ensure that AI/ML use is fit for purpose and shaped by the needs of the organisation, and avoid the risk of unnecessary and counterproductive change driven by the hype of AI transformation.

6.3 Directions for funder experimentation with AI

Finally, we briefly highlight future directions for funders to develop systematic experiments using the methodologies highlighted in our previous publication (Bendiscioli et al, 2022). Use of AI/ML should be evaluated in a similar way to any other process intervention, and both the work of funders and the development of AI/ML technologies will benefit from more systematic experimentation with AI/ML application.

We present here seven directions for potential experimentation by funders, beginning with AI applications closest to current use and going to more speculative directions for AI/ML use. Many of the directions highlighted here reflect directions explored in specific case studies in

Part 5, which can serve as a template for funders seeking to develop structured experimental evaluations.

6.3.1 AI in reviewer matching

Helping funders identify the best peer reviewers and panel members is currently the most common application of AI/ML in research funding (Rushforth et al, 2025). However, the contributions of AI/ML technologies to this process, and their impacts for the organisation, have not yet been systematically explored.

There is significant scope for exploration and experimentation in the types of AI/ML methods used: for example, different types of data used as input for AI/ML models (reviewer scientific record, application materials, applicant team information, etc); different modelling techniques; and different kinds of outputs and purposes for AI/ML application, such as a ranking of possible reviewers, or suggestions of new reviewers to add to an established reviewer pool.

For focused experimentation to compare and improve AI/ML approaches, funders could begin with curated reference sets of applications and “ideal” reviewer assignments, as selected by teams of scientific officers within the organisation. AI/ML systems could then be evaluated based on their ability to produce these ideal matchings, as an initial step in technical comparison.

To evaluate impact on the full review process, funders could develop a foundational A/B testing experimental framework to perform multiple experiments with different AI-enabled reviewer matching strategies. For example, funders could set up parallel experimental/control tracks within a single call to which applications are randomly assigned, and measure the impact of an AI-enabled matching system based on measures such as: time to completing reviewer recruitment; number of declined invitations; or scientific officer feedback on the quality of reviews received.

6.3.2 AI in peer reviewing

The use of AI/ML to help produce and process peer review reports has been explored in the experimental literature (Price & Flasch, 2017; Checco et al, 2021), but not yet systematically explored within funding organisations. Generative AI technologies have particular potential to help produce and summarise peer reviews, but also present significant risks of producing inaccurate and misleading content. Funders will benefit from developing structured evaluations of generative AI technologies in peer review processes.

Automated evaluation of AI-generated reviews is a complex process without clear measurement strategies (Yuan et al, 2022). Funders should therefore approach this in one of two ways: 1) expert evaluation of quality and informativeness of generated reviews, relying either on scientific officers or expert peer reviews; or 2) comparative evaluation of peer review processes, comparing manual peer reviews to those where AI is used as part of the reviewing process.

Both of these designs would enable experimentation with different AI-based generation or synthesis strategies, as well as different levels and types of AI use in the peer reviewing process. However, the reliance on expert evaluation will make these experiments time-consuming and expensive to run, and funders should therefore choose carefully what aspects they wish to prioritise in their experimentation.

6.3.3 AI for prioritising funding applications

Case study 5.3 illustrates one recent example of using AI/ML systems in the process of prioritising applications for consideration in the funding process. As an area where decisions have significant material consequences (i.e., affecting whether an application is awarded funding or rejected), a human-in-the-loop approach is essential. However, the ranking process is a clearly-defined problem for AI, and one where funders have developed clear criteria to guide the process. This means there is real potential for funders to benefit from careful experimentation with AI/ML to support the prioritisation process.

Automated evaluation for piloting different AI/ML systems is quite straightforward for funders to perform with their own historical data. Records of which applications were awarded and rejected, and what scores were assigned by panels provide data that can function directly for training and evaluating AI/ML systems for ranking applications. This provides a valuable and easily-implemented platform for testing and innovating with new AI/ML methodologies with no impact on new funding decisions.

Experimentation in live funding calls, however, requires a carefully structured approach with a well-developed plan in place for bringing the research community on board from the beginning of the process. A randomised controlled trial is a natural approach, with funding applications randomly assigned to a pool for manual or semi-automated prioritisation (with human oversight over all decisions), and evaluation based on final awarding decisions and time saved in the process. As a starting point, we refer funders who may consider this direction to the design implemented by “La Caixa” Foundation in their experimentation (Cortés et al, 2024).

6.3.4 AI in research assessment exercises

Case study 5.6 describes recent research evaluating the use of machine learning models to assist in automated scoring of research outputs in a national assessment exercise in the UK. A similar study has explored the use of machine learning in assessing research impact (Williams et al, 2023), and early analysis has examined the use of large language models for research assessment (Thelwall, 2024).

This is an area where the use of AI/ML techniques shows particular promise, but also significant challenges in how to effectively integrate the use of AI/ML without compromising the quality of assessment or the trust that governments and research communities place in assessment processes.

Funders should build on these initial examples to explore further targeted experimental questions in where and how AI/ML might best be leveraged within assessment processes. As there is already active academic research evaluating specific AI/ML approaches for assessment, funders will benefit from focusing particularly on evaluating which elements of research assessment processes are effective points for AI/ML intervention, and how to strengthen the trust built up with key stakeholders in assessment processes when AI/ML systems are introduced.

6.3.5 AI for applicant self-assessment

As well as their demonstrated value in helping to select submitted funding applications, AI/ML technologies have significant potential for applicants as tools to help self-assess and improve applications in development. This may come in the form of assistive writing technologies (e.g., dedicated writing assistants such as Grammarly, or more general-purpose generative AI platforms), or more purpose-built tools to score likelihood of success and generate feedback on in-progress application materials.

Self-assessment systems might therefore vary widely in terms of their generalisability across funders, but the most useful systems would be funder-specific. Automated evaluation of these systems would be possible to a limited extent by utilising previously collected applications and the peer reviews they received, but matching specific review feedback to particular aspects of an application is a challenge without good automated solutions for measurement.

Funder experimentation in this area would therefore best be designed with established user experience research methods such as focus groups and user studies. These studies could work with both applicants, to assess utility, and scientific officers, to assess how well aligned the feedback is with funder expectations. These types of evaluations would enable funders to compare different AI/ML strategies for self-assessment.

A more sophisticated evaluation of the impact of these systems on funding processes could be performed in two ways. Funders could use an opt-in approach to evaluate the outcomes of applicants who self-select for using a system compared to those who do not, or could provide a self-assessment system for a subset of funding calls and survey applicants to compare experiences where the tool was and was not made available.

6.3.6 AI for navigating funding resources

The use of AI/ML technologies is not limited to decision-making or assessment of quality. Funders also provide extensive information to applicants and policymakers, which can be difficult to navigate efficiently without prior experience. Extensive research on interactive AI systems to provide customer-facing information from large textual knowledge bases (Fader et al, 2014; Xu et al, 2024) is increasingly translating into everyday web applications for question answering which may be highly useful to help navigate funder resources.

Automated evaluation of interactive question answering systems is not straightforward, but there are strong precedents for developing curated sets of resources, questions to query them, and expected answers that can provide a starting point for funders to experiment internally with different AI/ML methods (Chen et al, 2019).

More direct experimentation with these types of AI/ML applications would best be performed by deploying AI systems as an optional tool on funder websites for users to interact with, and conducting opt-in surveys for users to report their experiences. A randomised controlled method is possible in this context, in which users are randomly assigned to a version of the website in which the AI tool is available or to the standard (non-AI) version of the funder website, and both populations are surveyed on their experience locating the resources they need.

6.3.7 AI in strategic planning

Strategic planning is a highly complex process, and each funding organisation approaches it in their own way. However, the value of AI/ML systems in supporting discovery and learning from large volumes of data suggests potential value in specific areas such as foresight activities or identifying directions for strategic funding calls. While experimentation with strategic planning is higher risk, identifying routes for success with AI/ML in this context can also be high-reward for funders willing to take on the experimental process.

To experiment with AI/ML in strategic planning, funders need to identify specific elements of their planning processes that may be amenable to AI/ML intervention. For example, landscape analysis can benefit from AI use to help identify emerging trends and patterns in current research (Holm et al, 2024); funders could also use machine learning analysis of applications and outcomes from past strategic funding calls to identify particular characteristics or directions the research communities they serve are especially responsive to.

Structured experimentation in strategic planning must be designed on a case-by-case basis. Funder research in this area is more likely to be analytic in nature, using AI/ML technologies to help learn about the research landscape or funding patterns, which can then support hypothesis-driven experimentation with the design of strategic funding calls.

6.4 Closing words

The use of artificial intelligence and machine learning in the work of research funding and assessment is an evolving area. Best practices and resources for funders exploring and applying AI/ML will continue to grow and change as AI technologies become more and more commonplace.

This handbook provides a snapshot of the current landscape of AI/ML use by research funders, and a starting point for funders to inform their future AI journeys. Drawing on two years of co-productive work and discussion with an international community of funders, we have presented here:

- ➔ Working definitions of responsible AI for funders
- ➔ The wider context in which funders explore and apply AI/ML
- ➔ Practical steps involved in developing AI/ML applications in funding organisation
- ➔ Organisational issues and strategies informing the use of AI/ML
- ➔ Case studies of diverse real-world applications of AI/ML by research funders