HIGH5 for AI Supplementary Report: 5.2 Hackathon Proposal

Report number OR/26/050_subreport5

Authors:

Lucy Killoran [1] https://orcid.org/0000-0002-1907-6078
Jamine Hunter [3] https://orcid.org/0000-0001-8643-2602
Jesse Alexander [1] https://orcid.org/0009-0006-2877-3197
Carl Watson [2] https://orcid.org/0000-0003-4940-5503

Affiliations:

1 - Science and Technology Facilities Council
2 - British Geological Survey
3 - UK Centre for Ecology and Hydrology

Executive Summary

The HIGH5 for AI project set out to increase the availability and accessibility of high-value environmental datasets for users beyond the traditional environmental science community. This report, produced as part of Work Package 5, presents a practical framework for using hackathon-style events as a mechanism for achieving that goal. It explores what form those events should take, what they should focus on, and how they should be run.

Drawing on user research from across the HIGH5 project, AMPLIFY-EDS, a review of relevant literature, and market research interviews, the report reaches a clear conclusion: **hackathons can deliver genuine value, but only when they are rigorously designed, adequately resourced, and built around challenges that reflect real end-user needs**. Poorly scoped challenges, under-preparation, and unclear outcomes are the primary causes of failure. The evidence from both case studies is unambiguous: preparation is not optional, and the investment required to run a successful event should not be underestimated by organisers or funders.

This report frames the most important factors to consider when designing a hackathon through three 'Tasks'. Task 1 is to define the purpose of a hackathon through exploring a range of potential outcomes and challenges. Task 2 is to conduct market research to understand the key success factors, risks, and costs of hackathon-style events. Task 3 is to integrate research insights into a full strategic and operational event plan. This report documents the completion of these tasks, with concrete ideas, suggestions, and recommendations made throughout.

Importantly, hackathons are only part of the picture. **To achieve lasting impact, hackathon events must be embedded within a broader programme of stakeholder engagement and community support**. Attracting participants from outside the environmental science community, sustaining their interest beyond the event itself, and translating hackathon outputs into real-world adoption all require ongoing effort that goes well beyond what any single event can deliver. Hackathons are most effective as catalysts—accelerating progress and building relationships—when they are supported by wider outreach, accessible onboarding materials, and sustained community development activity.

The report concludes with four recommendations for funders. Firstly, a **one-day ideation and consensus-building workshop** should be a priority to ensure the success of a HIGH5 hackathon. This event would bring together HIGH5 tool and resource developers, senior EDS decision-makers, and staff responsible for running the proposed event(s) to develop candidate hackathon designs for further development. The second and third recommendations are an either/or choice: to run a programme of hackathon-based events to **engage diverse audiences**; OR run a one-off hackathon event to **produce targeted outcomes**. The main takeaway from this report should be that a key success factor of hackathons lies in their specificity. This is reflected in these second and third recommendations, which provide two approaches to balancing event specificity with audience engagement. The final recommendation is that the practical lessons from this work should be distilled into **guidance for the EDS Co-Creation Toolkit**, ensuring that the investment made here continues to benefit future event organisers across the environmental data community and beyond.

Contents

About the report 4

Task 1: Defining outcomes and challenges 5

What is a hackathon? 5

What is the purpose of a HIGH5 hackathon? 5

What topics does prior work indicate users want to explore? 5

What HIGH5 deliverables are available for use in a hackathon? 7

Potential outcomes and challenges for a HIGH5 hackathon 7

Task 2: Market research 11

Summary of findings 11

Case study 1: A data-centric hackathon run by a data science organisation 12

Purpose 12

Key success factors 12

Risk mitigation 12

Costs and resourcing 12

Case study 2: A product-centric hackathon run for a domain-specific organisation 13

Purpose 13

Key success factors 13

Risk mitigation 13

Costs and resourcing 13

Task 3: Bringing it all together: A hackathon design tool 14

A hackathon design tool for the EDS 14

Three models for a HIGH5 hackathon 15

H1: Attached to environmental science community event 15

H2: Engaging industry and startups 16

H3: Thematically aligned with government mission 17

Recommendations and next steps 18

Supporting materials 20

Appendix A: Definition of terms (after Kollwitz and Dinter, 2019) 21

About the report

This report will be useful to potential organisers and funders of hackathon-style events.

About the Environmental Data Service

The Environmental Data Service (EDS) provides a focal point for scientific data and information spanning all environmental science domains: atmosphere and climate, earth observation, polar and cryosphere, marine, terrestrial and freshwater, geoscience, solar and space physics. The EDS is made up of a network of distributed data centres, with domain specific expertise.

Our main goal is to ensure that environmental data are made available, accessible and re-usable for the long-term in order to fully realise their value. We are funded by the Natural Environment Research Council (NERC), part of UK Research and Innovation, to advise researchers on how to prepare data for long term storage and dissemination.

The EDS is a fundamental part of NERC's digital strategy and works with the Digital Solutions Hub, and other partners, to break down disciplinary barriers and facilitate data sharing beyond academic use.

About the project

The NERC funded HIGH5 for AI project (hereafter referred to as HIGH5) aims toto increase the availability of high-value AI-ready EDS datasets through the development of APIs and AI tools for the environmental science community and beyond. The project, undertaken between October 2025 to March 2026, was divided into five key work packages:

- **WP1** focused on project management and communication
- **WP2** delivered internal and external co-design activities
- **WP3** implemented the data APIs themselves
- **WP4** developed an Agentic AI capability demonstrator
- **WP5** focused on enabling users beyond the traditional environmental data community

This report aims to summarise the work undertaken in WP5.2, which focused on analysing what would make a successful HIGH5-related hackathon and developing options for such events. This report presents the following contribution towards this goal:

- Summarising the objectives of hackathon-style events, based on current understanding of our stakeholder needs
- Exploring best practices, as reported by literature and hosts of past events
- Outlining a framework for planning hackathon-style events, with a focus on addressing user centred challenges within an environmental data service context
- Identifying opportunities and recommendations for the short- and medium-term future, both within environmental data services and more broadly, to support open science, encourage collaboration and improve data accessibility for cross-disciplinary users

Task 1: Defining outcomes and challenges

What is a hackathon?

Hackathons are events where mixed groups work collaboratively on problems posed as 'challenges' in a time-constrained setting. Hackathon challenges traditionally focus on building prototype software solutions or undertaking a piece of data analysis (the latter often being called datathons). The hackathon-style format has also been applied outwith these traditional contexts, with challenges and outcomes customised to meet the needs of specific communities. To reflect this versatility, hackathons can be broadly described as events which:

> "[A]ccelerate the development of scientific results and collaborations, communities, and innovative prototypes [to] address urgent challenges" (Falk et al., 2024).

Two key tasks for HIGH5 are therefore to determine the **types of outcomes** which are to be accelerated and the **nature of the challenges** that the event would be designed around

What is the purpose of a HIGH5 hackathon?

HIGH5's project-level goals include promoting the tools developed to users within and outside of the EDS; the wider aspiration being to enable the EDS to deliver value beyond the environmental science community by engaging with a broad range of disciplines and sectors and addressing key government missions. Key government missions include clean energy and a future-proof NHS, which HIGH5 responds to through its focus on two example use cases of offshore wind farms and blue-green spaces.

> To represent value for money, the emphasis of a hackathon must be on the accelerated outcomes and the challenges which will drive the event

Two important factors should be noted when considering the purpose of a HIGH5 hackathon, which are the **emphasis of hackathons on challenges**, and the **investment required** to run a hackathon-style event. Although a hackathon presents the opportunity to showcase the tools developed, gather evidence on how far they meet known requirements, and generate new insights to guide continued development, the significant resources required to run a hackathon-style event means that an event focused only on user evaluation and feedback would likely not represent a good return on investment. Focused user research towards these aims could be completed to a high standard using significantly less resources. To represent value for money, the emphasis of a hackathon must be on the accelerated outcomes and the challenges which will drive the event.

What topics does prior work indicate users want to explore?

User research recommendations from HIGH5 WPs 2 and 5 and AMPLIFY provide a starting point for thinking about the purpose, outcomes, and challenges of a HIGH5 hackathon.

High-level recommendations from user research show that there is a clear need for access to multi-domain environmental data, but that this is hindered by data sources **scattered across multiple locations**, and a level of **standardisation which does not meet requirements**. Users frequently request solutions for improved data interoperability through products such as pre-aggregated data in thematic packages. Under these conditions, users have expressed that their primary channels for finding high-value data are their **personal networks and institutional relationships**. However, users acknowledge that this is not ideal and express the desire for a 'single source of truth'. User reliance on personal networks and institutional relationships rather than data archives may be a result of their main pain point: **time scarcity**.

Increased FAIRness of EDS data would reduce the time and effort required by users to find, verify, download, and use data. For example, users have requested improved support for specific interactions related to finding, accessing, interoperating, and reusing data, such as:

- **Data discovery**, e.g.
    - High-quality metadata
    - Consistent terminology
    - Standardised interoperable formats
    - Clear signposting to high-value data
- **Dataset value verification before download**, e.g.
    - Preliminary data checking, analysis, and visualisation
    - Interactions which support the visual exploration of data
- **Customisable download options**, e.g.
    - API endpoints
    - Subsetting of large datasets
    - Processed products, such as 'analysis ready' data and pre-aggregated thematic packages

Standardisation and metadata are key components of increased FAIRness. Users have shown interest in approaches to standardisation which can **flexibly harmonise** data across multiple sources, as opposed to those which introduce additional rigid or static standards.

Crucially, users also want **plain-English advice** to support their interactions with EDS services. This suggests a need for improved onboarding as well as support for ongoing service use. Alongside improved documentation, users perceive AI-based tools to offer potential value here and can imagine using natural-language AI tools to assist them in discovering, verifying, and downloading data.

This research suggests that EDS users may be interested in exploring the following broad topics in a hackathon environment:

- Approaches to flexibly harmonising data and metadata across multiple sources

- New types of interaction for data discovery and verification before download, including AI assistance
- APIs for data access, subsetting, and download
- Data products such as analysis-ready data or thematic data packages

Importantly, these needs will vary depending on user motivations. Three types of EDS service user were developed in AMPLIFY, which can be used to give a sense of the motivations and interests of prospective hackathon participants:

- **Technology and software professionals**
  - Responsible for technical creation and integration of digital solutions
  - Includes software developers focused on the digital solution rather than producing insights from the data
  - API access is critical for their daily work
- **Data analysis and integration professionals**
  - Require programmatic access (APIs) to combine, manage, and analyse large or diverse environmental datasets
  - Includes GIS Specialists and a growing cohort of researcher software engineers (RSEs) focused on discovering insights from the data
- **Industry and infrastructure professionals**
  - Require automated integration with environmental data to assess risks and operational performance of a system
  - Includes staff within utility companies and regulators

What HIGH5 deliverables are available for use in a hackathon?

HIGH5 directly responds to many of the recommendations made in prior user research by developing **API and chatbot interfaces** which expose **high-value EDS data**. The development of APIs aligns with accessibility requirements, and this work also involves exploring the use of **Croissant as a metadata standard** for machine-learning-related accessibility and interoperability. The APIs and Croissant records expose analysis-ready versions of high-value EDS datasets, improving data reusability. Findability of data and services is supported by the development of targeted **onboarding materials** and chatbot interactions focused on data discovery. The potential real-world applications of these tools are presented through HIGH5's two example use cases of blue-green space and offshore wind farms.

It is crucial that these resources are **assessed for readiness** before any user-facing event is designed around them. Aspects to consider as part of readiness checks are presented in the market research section of this report.

Potential outcomes and challenges for a HIGH5 hackathon

Table 1 synthesises the findings of prior research with the deliverables of HIGH5 to give a clear overview of the **points of focus** that a HIGH5 hackathon could be designed around. Table 2 presents three possible combinations of these points of focus to show how these elements can work to **construct a hackathon challenge**. Table 2 presents only three of many possible combinations of Table 1's points of focus, underlining the thought required in designing a **meaningful hackathon challenge** and the **highly contrasting outcomes** which can result from different combinations.

Category	Points of focus
HIGH5 tools	• APIs • High-value, analysis-ready datasets • Croissant records / croissant as metadata standard • Agentic AI chatbot
Topics of interest	• Approaches to flexibly harmonising data and metadata across multiple sources • New types of interaction for data discovery and verification before download, including AI assistance • APIs for data access, subsetting, and download • Analysis-ready data or thematic data products
User types	• Technology and software professionals • Data analysis and integration professionals • Industry and infrastructure professionals
Government mission themes	• Public health, an NHS fit for the future • UK clean energy provision
Engagement with broad range of disciplines and sectors	• Urban blue-green space stakeholders • Offshore wind farm stakeholders
Outcomes to be accelerated	**Umbrella outcomes:** • Engagement of disciplines and sectors outwith environmental science • Application of HIGH5 tools to real-world challenges **Topic-specific outcomes:** • Software prototypes (e.g. package, app, dashboard) • Dataset products (e.g. analysis-ready, thematic) • ML models (e.g. classifier, detector, recommender, alerter) • Interaction guidelines (e.g. responsible AI for health data discovery, API documentation) • Proposals or roadmap for future development of tools • Domain-specific or cross-domain use cases

Table 1: An overview of recommendations for the focus of a HIGH5 hackathon, based on a synthesis of HIGH5 deliverables and findings arising from prior research.

Challenge	Tool	Topic	User type	Theme	Sectors
Use environmental data APIs to build tools that help health professionals, urban planners, and policymakers' access and act on environmental data—without needing specialist knowledge of environmental science	APIs	APIs for data access, sub-setting, and download	Technology & software	Public health	Health Software UX/HCI[1] Data
Build AI-assisted interfaces that let energy developers, engineers, and regulators find, assess, and download the environmental data they need—reducing reliance on specialist intermediaries	AI chatbot	AI assistance for data discovery and verification	Industry & infrastructure	Clean energy	Energy Software UX/HCI AI Data
Use analysis-ready environmental datasets to generate insights that are directly usable by public health researchers, clinicians, or local authorities who wouldn't typically work with environmental data archives	High-value datasets	Analysis-ready / thematic data products	Data analysis & integration	Public health	Health Data ML

Table 2: Example hackathon challenges constructed from three possible combinations of Table 1's points of focus. Different combinations can produce highly contrasting outcomes, underlining the thought required in designing a meaningful hackathon challenge. The challenges presented in this table are suggestions based on the work presented in this report so far. ***The final challenge for any HIGH5 hackathon should be designed collaboratively by senior EDS decision-makers and the staff responsible for delivering the event(s).***

[1] UX/HCI: User Experience / Human-Computer Interaction

Task 2: Market research

This section reports on two market research interviews which were conducted with organisations who have previously run hackathon-style events. The interview findings are presented as two case studies, each describing the event's **purpose**, **key success factors**, **risk-mitigation advice**, and **costs**.

Summary of findings

The findings of these case studies emphasise that hackathon-style events can deliver high value for skills development and exploratory research, but they **should not be relied upon as primary mechanisms for producing deployable technical solutions**. Their success depends overwhelmingly on rigorous preparation, clear challenge scoping, and substantial staff and logistical resourcing. When poorly designed or under-resourced, they carry high risks of inefficiency, participant burnout, and limited usability of outputs. Key risks and their mitigations are outlined in Table 3.

Preparation requirements include **rigorous assessment** of the appropriateness, reliability, sensitivity, and sufficiency **of provided resources**, and **feasibility of challenges**; along with potential involvement of legal and IP teams. Lack of upfront preparation shifts pressure into the event and reduces outcomes. During the event, **clear ownership** (e.g. designated challenge leads) and **continuous access** to domain experts, technical support, and facilitation improves coherence and quality. Momentum and morale are also directly affected by **team size and composition**.

Key risk	Negative impacts	Mitigation
Unclear or over-flexible challenges	Duplicated, impractical, or low-impact outputs	Ensure challenges are tightly defined with explicit constraints
Misaligned expectations	Disappointment if polished solutions are expected	Explicitly frame outcomes as exploratory, or serving goals such as engagement or collaboration
Competition-based judging	Damaged morale if perceived as rewarding 'style over substance'	Consider collaborative or reflective wrap-ups instead
Underestimated staff time	Poor quality event with negative reputational impacts	Plan for substantial staff involvement before, during, and after the event
Underestimated costs	Poor quality event with negative reputational impacts	Funding which supports a clear development, training, collaborative, or exploratory objective

Table 3: A summary of the key risks, their potential negative impacts, and advice for mitigations from the market research case studies.

Case study 1: A data-centric hackathon run by a data science organisation

Purpose

A **non-competitive, collaborative research hackathon** focused on peer-to-peer learning and exploratory data science. This style of event is deliberately not prize-driven and does not guarantee a solved problem. The main outcome is a public final report, often serving as a springboard for longer-term projects. The main recommendations are **extensive challenge preparation** with challenge owners, and **structured onboarding**.

Key success factors

- Extensive challenge preparation is critical and can take 5 months to >1 year—poorly scoped challenges are the biggest risk to failure
- Challenges must be feasible in five days, using mostly established methods, and driven by what the datasets can realistically support
- Structured onboarding (a full precursor week) where participants can ask questions of challenge owners means immediate productivity during the main event
- Teams of 10–12 participants work best with a trained facilitator and explicit guidance on teamwork and documentation
- Employing an early-career researcher as PI is central: they scope the challenge, support teams, and write the final report, ensuring coherence and academic quality
- Strong logistical and technical support (IT, compute, facilitation, catering) reduces friction and burnout during long working days

Risk mitigation

- Rigorous data readiness assessment (i.e. appropriateness, reliability, sensitivity, and sufficiency)
- Clear expectation-setting: exploratory outcomes, not polished products
- Early involvement of legal teams (IP, data sensitivity, participant agreements)

Costs and resourcing

- Travel, accommodation, and catering for participants
- Compute resources (~£3–3.5k Azure credits per event)
- Social event (~£2k for 30–40 people)
- Significant staff time (project officers, PhD support, IT, PI effort)
- Economies of scale apply, i.e. venues in major cities may be more expensive but travel may be cheaper than to less well-connected locations, and vice versa

Case study 2: A product-centric hackathon run for a domain-specific organisation

Purpose

This hackathon focused on **early-career participants** using domain-specific datasets to prototype digital products. Primary benefits were **skills development**, **engagement**, **and organisational insight**, rather than deployable tools. A major lesson was that unclear outcomes led to limited long-term impact: none of the prototypes were developed further.

Key success factors

- Worked well as learning and employability events, building technical, teamwork, and presentation skills, but less as a way to develop useable technical solutions
- Small teams (4–5 people) with deliberate skills and gender balance improved dynamics of teamwork
- Continuous participant support (domain experts, data access help, wellbeing checks) was essential due to the intensity of the environment
- In-house catering and a well-designed venue prevented disruption and loss of momentum

Risk mitigation

- Vague or overly flexible challenges led to duplicated or impractical outcomes—clear problem definitions are strongly recommended
- Judged competitions can undermine morale if winners are perceived as “shiny but unhelpful”; alternative wrap-ups may be better for smaller events
- Lack of advance participant preparation increased pressure during the event
- Hackathons are resource-intensive; underestimating staff time is a major risk
- Hackathons should be used selectively, with clear goals and realistic expectations, especially in constrained funding environments

Costs and resourcing

- Fully funded by an external academic body as primarily served as a career development opportunity for early-career researchers
- Specialist external organisers reduced risk but still required significant staff involvement and time which is not to be underestimated
- Major costs included venue, catering, materials, facilitation, and staff mentoring time
- Financial viability would have been challenging if funded through another route, e.g. paid places were hard to fill

Task 3: Bringing it all together: A hackathon design tool

A hackathon design tool for the EDS

The sections above synthesise research and HIGH5 deliverables to suggest possible outcomes and challenges for a HIGH5 hackathon, as well as the key risks and mitigations to consider. This section presents an approach which can be used to bring these insights together into a plan for an event which is customised to the needs of the EDS. This tool is a taxonomy of the essential components of any hackathon-style event (after Kollwitz and Dinter, 2019)—presented below, in Table 4. This tool breaks a hackathon down into 11 **dimensions**, each with 2–3 **characteristics**. These dimensions are categorised by whether they fall into **strategic** or **operational** aspects of the event (i.e. the nature of the challenge and outcomes versus the practical logistics). Characteristics from each dimension can be mixed and matched to develop a full strategic and operational plan for a hackathon.

	Dimension	Characteristic		
Strategic (nature of the challenge)	idea stage	generation	conversion	diffusion
	challenge type	technology- centric	topic-centric	data-centric
	solution space	open	semi-structured	structured
	outcome maturity	broad concepts	conceptual solutions	functional solutions
	value proposition	challenge output		participant benefit
Operational (practical logistics)	duration	~1 day	1–3 days	>3 days
	venue	virtual	hybrid	in-person
	incentives	competition		collaboration
	cost to participant	no cost	some cost	full cost
	target audience	public	research	industry
	resources	not provided	partially provided	provided

Table 4: After Kollwitz and Dinter (2019), a taxonomy of hackathon characteristics which can be used as a tool for designing hackathons. Characteristics from each dimension can be mixed and matched to develop a full strategic and operational event plan. See Appendix A for definitions of the more ambiguous terms.

The following sections illustrate this tool in-use through three model events. **NOTE: These event models are independent of the ideated challenges presented in Table 2.** Each of the below models describes a different approach to running and resourcing an event. The final design of any event(s) should be driven by intentional choices regarding target audience, available funding, and desired

outcomes, with the specifics of the challenge designed collaboratively by senior EDS decision makers and the staff responsible for delivering the event(s).

Three models for a HIGH5 hackathon

H1: Attached to environmental science community event

	Dimension	Characteristic		
Strategic	idea stage	generation	**conversion**	diffusion
	challenge type	**technology- centric**	topic-centric	data-centric
	solution space	open	semi-structured	**structured**
	outcome maturity	broad concepts	**conceptual solutions**	functional solutions
	value proposition	**challenge output**		participant benefit
Operational	duration	**~1 day**	1–3 days	>3 days
	venue	virtual	hybrid	**in-person**
	incentives	competition		**collaboration**
	cost to participant	no cost	some cost	**full cost**
	target audience	public	**research**	**industry**
	resources	not provided	partially provided	**provided**

Table 5: An event model for a hackathon attached to environmental science community event. See Appendix A for definitions of terms.

As part of an environmental science community event, this one-day, in-person hackathon brings together researchers and industry practitioners to collaboratively convert promising ideas into clearly defined concepts. A **technology-centric challenge** operates within a structured solution space, with predefined tools, data, and technical requirements provided to participants. Teams focus on producing **conceptual, non-functional solutions** such as mock-ups, models, or system designs that directly respond to the challenge brief. With collaboration as the primary incentive and full participation costs covered by attendees, the hackathon emphasizes the creation of tangible challenge outputs that can **inform future research–industry partnerships** and **downstream innovation** in environmental science.

H2: Engaging industry and startups

	Dimension	Characteristic		
Strategic	idea stage	generation	**conversion**	diffusion
	challenge type	**technology- centric**	topic-centric	data-centric
	solution space	open	semi-structured	**structured**
	outcome maturity	broad concepts	conceptual solutions	**functional solutions**
	value proposition	**challenge output**		participant benefit
Operational	duration	~1 day	1–3 days	**>3 days**
	venue	**virtual**	hybrid	in-person
	incentives	**competition**		collaboration
	cost to participant	**no cost**	some cost	full cost
	target audience	public	research	**industry**
	resources	not provided	partially provided	**provided**

Table 6: An event model for a hackathon which engages industry and startups. See Appendix A for definitions of terms.

This virtual, multi-day hackathon brings together industry teams and innovative startups to convert promising early ideas into viable solutions. Participants tackle a **technology-centric challenge** within a highly structured solution space, using provided resources such as data, APIs, tools, and expert support. Over more than three days, teams collaboratively refine selected ideas into **functional, proof-of-concept solutions** that directly address the defined challenge. The hackathon operates as a free-to-participate, virtual, competitive event, encouraging focused problem-solving and performance. Its primary value lies in the challenge outputs: practical solutions designed to **meet industry needs** and **spark further development or adoption**. It is strongly advised that organisations seek legal advice on participant agreements for this style of hackathon regarding intellectual property rights to any data, concepts, or products provided for or developed during the event.

H3: Thematically aligned with government mission

	Dimension	Characteristic		
Strategic	idea stage	generation	**conversion**	diffusion
	challenge type	technology- centric	**topic-centric**	data-centric
	solution space	open	**semi-structured**	structured
	outcome maturity	broad concepts	**conceptual solutions**	functional solutions
	value proposition	**challenge output**		participant benefit
Operational	duration	~1 day	**1–3 days**	>3 days
	venue	virtual	hybrid	**in-person**
	incentives	competition		**collaboration**
	cost to participant	no cost	**some cost**	full cost
	target audience	public	**research**	industry
	resources	not provided	partially provided	**provided**

Table 7: An event model for a hackathon which is thematically aligned with a government mission. See Appendix A for definitions of terms.

This 1–3 day, in-person hackathon supports a government public health mission by helping research participants convert existing ideas into promising concepts. Framed around **a challenge on the topic of public health**, teams work within a semi-structured solution space; they are guided by predefined datasets, policy goals, and technical requirements while retaining some freedom to integrate their own ideas. Participants develop **conceptual, non-functional solutions** such as mock-ups, models, or frameworks. The event targets a cross-disciplinary research audience, so does not develop industry- or public-facing angles on the challenge. It involves some participant cost, so emphasises the value of **shared resources**, **expert support**, and **collaboration** over competition to **foster enduring research partnerships** based on further developing the conceptual outcomes.

Recommendations and next steps

This report demonstrates that a well-designed hackathon represents a significant opportunity for the EDS to extend the reach and impact of HIGH5 tools and datasets beyond the environmental science community. However, it has also shown that hackathons are resource-intensive and must be carefully designed to deliver return on investment.

For instance, this report has explained why hackathons can deliver high value for outcomes such as skills development and exploratory research but should not be relied upon as primary mechanisms for producing deployable technical solutions. Additionally, poor quality events may not only fail to deliver desired outcomes but may also have negative reputational impacts—producing the opposite of the cross-sector engagement desired. Because of the value of these findings, the first recommendation is that a version of this report is written up for the EDS Co-Creation Toolkit:

R1	**A write-up of this report for the EDS Co-Creation Toolkit** The practical lessons learned through this project about designing hackathons for environmental data contexts should be written up as a short, accessible blog-style post for the EDS Co-Creation Toolkit, linking through to this report for readers who want the full detail. This will ensure the investment made in this work continues to benefit future event organisers across the EDS community and beyond.

The remaining recommendations relate to running HIGH5-related hackathon events. We recommend that funders consider supporting the below actions. Note that R3 and R4 are either/or options—the outcome of R2 would be adjusted depending on whether R3 or R4 was chosen to follow.

R2	**A one-day ideation and consensus-building workshop** This workshop should be considered a priority for ensuring the success of any HIGH5 hackathon. This event would bring together HIGH5 tool and resource developers, senior EDS decision-makers, and staff who would be responsible for running the proposed event(s) to further explore the content outlined in Task 1 and Task 3 of this report. This workshop would support stakeholders to test the readiness of tools and resources; ideate outcomes, challenges, and model events; and build consensus around one or more candidate events to develop further. This workshop may highlight the need for an additional period of development to improve the readiness of HIGH5 tools and resources for user-facing events.

AND

R3	**A programme of hackathon-based events to engage diverse audiences** The three event models outlined above offer a realistic and scalable pathway from low-risk, low-cost first steps through to high-impact flagship events. Building on the outcomes of the R1 workshop, a series of hackathon-style events would form a programme of stakeholder engagement, each building confidence and evidence for the next. For example:

	1. Run an H1 event aligned with the NERC Digital Gathering 2026 The NERC Digital Gathering provides a ready-made audience and reduces organisational overhead, making it the lowest-risk starting point. However, to realise the broader ambitions of HIGH5, deliberate effort must be made to attract participants from outside the traditional UK environmental research community, including software developers, data engineers, and industry practitioners. This means active outreach beyond standard NERC channels well in advance of the event. **2. Run an H2 event in collaboration with Geovation** Geovation's existing network of geospatial startups and industry partners offers a direct route to audiences who would benefit from HIGH5 tools but are currently underserved by EDS. A virtual, competitive format keeps costs manageable while maximising reach. This collaboration should be initiated promptly to take advantage of current momentum from the HIGH5 project. **3. Subject to funding, run a residential H3 event focused on health and environment** This is the most ambitious and highest-impact of the three models. A residential format, centred on the public health and environment theme, would bring together cross-disciplinary researchers to develop conceptual solutions with real policy relevance. Critically, this event should be accompanied by a professional communications effort—working with communications experts to tell the story of health and environment interactions to a wider public audience, extending impact well beyond the event itself. The programme does not need to follow these event models; this is an illustration to show how diverse audiences could be engaged through different event designs.

OR

R4	**A one-off hackathon event, intentionally designed to produce targeted outcomes** Building on the outcomes of the R1 workshop, one event would be used to extend the reach and impact of HIGH5 tools and datasets beyond the traditional environmental science community. It should be expected that the R1 workshop will produce multiple valuable candidate events; however, the ultimate success of a one-off event depends on an intentional design which can feasibly achieve specific outcomes. The temptation to combine multiple candidate ideas into one event should be avoided. Event designs which combine multiple topics, themes, challenges, or outcomes should be assessed against the risks outlined in this report

Supporting materials

Falk, J. et al. (2024) 'The Future of Hackathon Research and Practice', IEEE Access, 12, pp. 133406–133425. Available at: https://doi.org/10.1109/ACCESS.2024.3455092

Kollwitz, C. and Dinter, B. (2019) 'What the Hack? – Towards a Taxonomy of Hackathons', in T. Hildebrandt et al. (eds) Business Process Management. Cham: Springer International Publishing, pp. 354–369. Available at: https://doi.org/10.1007/978-3-030-26619-6_23

Appendix A: Definition of terms (after Kollwitz and Dinter, 2019)

idea stage	generation	generate initial innovative ideas
	conversion	select promising ideas for further development
	diffusion	test existing tools, products, and services
challenge type	technology- centric	software, hardware, or API-related
	topic-centric	solving a specific research or business problem
	data-centric	processing, analysis, or visualization of data; the collection or generation of data
solution space	open	wide-ranging challenges that leave plenty of room for interpretation, with minimal requirements and restrictions that could potentially limit creativity
	semi-structured	certain specifications limit the solution space, e.g. specification of technologies, data, methods, technical and/or functional requirements, with room left for individual approaches
	structured	strict demands on the procedure and results, which significantly limits the solution space
outcome maturity	broad concepts	operates at a high level of abstraction with a focus on creativity
	conceptual solutions	demonstrative, non-functional, paper-based or computer-aided mock-ups that represent a concept
	functional solutions	includes core functions and demonstrates the general feasibility of a concept (proof of concept), or represents a minimum viable product
value proposition	challenge output	utilises the resulting outcomes, e.g. ideas, models, prototypes, data visualisations; extension, improvement, or evaluation of existing entities
	participant benefit	generates benefits with respect to the participants, e.g. collaborations, educational, networking, career
resources	resources made available to the participants before or during the hackathon e.g. hardware, software, data; existing ideas, concepts, or prototypes to be developed further; human resources such as mentors or experts – may be voluntary or mandatory depending on solution space	