The experiment begins: Arcadia publishing 1.0

Building on the open-source platform PubPub, we're sharing the first iteration of our publishing website. In addition to posting our first set of research pubs, we're documenting our progress in developing this new system for sharing science and hope you'll provide feedback.

Version 3, published May 9, 2023. Originally published May 30, 2022.

Arcadia Science DOI: 10.57844/arcadia-050a-q254

Purpose

In thinking about how to share Arcadia's research, we wanted to keep features of traditional publishing that have been honed over centuries, but improve upon what hasn't quite adapted to the nature of modern science and technology. We have a unique opportunity to use our own research to develop mechanisms of sharing and quality control that can be more agile and adaptable. Our initial attempt is outlined here and we will continue to iterate upon it, always keeping the advancement of knowledge as our guiding principle when making decisions on what to try next.

This pub is intended to help you understand our thinking thus far, to provide a sense of what we've done and how the platform works, and to serve as a place to provide feedback on our strategy and the platform itself.

- This pub is part of the project, "Reimagining scientific publishing." Visit the project narrative for more background and context.

Introduction

We are reimagining scientific publishing — sharing our work early and often, maximizing utility and reusability, and improving our science on the basis of public feedback.

This is our first draft. We have ambitious goals and we're committed to replicable long-term solutions, but we also know that "perfection is the enemy of good." We're using this platform to release findings now rather than hiding them until we've gotten everything exactly how we want it. Readers can think of the pubs on this platform as drafts that will evolve over time, shaped by public feedback. The same goes for the platform itself! We're treating our publishing project like an experiment — we're not sure where we will land, but we can only learn if we try. In this pub, we're sharing our strategy and the reasoning behind some of our key decisions, highlighting features we're excited about and areas for improvement.

We don't want to clutter our scientific pubs with comments about website functionality, aesthetics, or our model, but we invite you to share those thoughts here! All first drafts need feedback. What do you think of this site? Are our first pubs understandable? What important features are we missing? What are we doing right? What can we improve? Please don't hesitate to share constructive commentary and ideas.

The platform

When starting a new experiment, it's tempting to build every piece from scratch. We could try to build an ideal publishing mechanism where form seamlessly fits function. However, we are a research organization with the goal of advancing our science with the maximal impact on the rest of the scientific community. If we held back our science from the world so we could build a platform from scratch, we would be fundamentally obstructing our experiment from its main purpose.

Enter Knowledge Futures Group (KFG) and PubPub. While there are features we intend to expand and customize, PubPub meets our main requirements for functionality and, importantly, is run by a team that is deeply aligned with our

open science mission. KFG is already developing functionality to make scientific research more mineable and its usage quantifiable. Their forward-thinking plans were not based exclusively on meeting today's publishing needs, but preparing for a world in which there is demand for exactly what Arcadia wants: an author-driven, maximally reusable, and community feedback-strengthened research-sharing platform. We hope that others can try our approach. Ultimately, the true measure of our success will be other scientists adopting, and adapting, publication paradigms that better serve science.

Initial framework

This section is a summary of our publishing strategy and what we've decided to try initially. Subsequent sections expand upon some of the more complex points. Check out the table of contents (click "CONTENTS" on the right side of the screen!) to get an overall sense of how this pub is organized, and check out whichever topics are most interesting to you.

Where can our research be found?

- Raw data will go in FAIR (Findable, Accessible, Interoperable, Reusable) repositories.
- Detailed protocols will be deposited in protocols.io.
- Contextual information for the raw data will live on the PubPub platform.

How can the scientific community use our work?

- Each pub is citable and will ultimately have a permanent identifier (like a digital object identifier, or DOI, through DataCite).
- Each pub will follow rules for Google Scholar indexing and be discoverable.
- Each pub will follow whatever structure and include whatever media can best convey the desired information to the intended audience. This may mean including executable code blocks, interactive content, and other resources embedded within pubs.

What do we intend to share?

- The nature of each pub will depend largely on the kind of data being reported and be geared towards facilitating use/reuse.
- We will include information about planned experiments/approaches, informative failures, potential uses by others, explicit questions soliciting feedback, and other components not typically included in scientific literature.
- An integral part of every research product will be public feedback that will be visible alongside the original pub. We will use this to iterate and improve on the work.
- We will provide context through “project narratives,” running logs that link individual pubs and keep the big picture in focus.
- While versions of the work will have permanent identifiers so they can be cited, work is not static — evaluation will happen after the pub is publicly released. The products will be improved/built upon over time, incorporating feedback throughout.

Who will contribute to our pubs?

- Each person who contributed to different aspects of the science will be listed as a contributor — including those commenters whose critical feedback shaped the work.
- Each contributor’s role(s) will be stated explicitly and in a standardized format, making the nature of their involvement clear.
- A subset of contributors responsible for the work will be “byline” contributors. All contributors will be listed in the metadata for the DOI, but only byline contributors will technically be listed as “authors,” and have their names used in citations.
- Arcadia’s byline contributors will be empowered to release their research with the support of our publishing team as needed. Wherever possible, we want to facilitate sharing rather than restrict or delay it. Anyone who makes a PubPub account can add comments on our pubs. We will also capture as much of the Twitter conversation as possible by embedding a feed collecting mentions of each pub.

When will we share our work?

- We will post publicly at multiple stages of a project. Pubs are modular and we will share them when we’re ready for community feedback or think we

have information that could be useful to others rather than stuffing more and more information into a single publication [1].

- We will improve pubs over time as we receive feedback and do more work. For transparency, each version of the pub will remain available.
- There will be synthesis works posted as well, which integrates modular data already released. This will be accomplished partially through project narratives, which track the running progress of a project as whole by stringing together individual pubs. In cases where we want to elaborate, we will share separate review/perspective pubs.

Site organization

As we create more content, the structure will likely shift and expand. For now, there are three layers to this site:

1. The **landing page** provides background on Arcadia and our goals, and, most importantly, will tell you what's new. [View landing page]
2. **Project narratives** are running logs that maintain a high-level view of our goals and connect modular pubs together into a more cohesive story. [View example]
3. **Pubs** are detailed research products. These are akin to traditional papers, but tend to have a narrower focus and looser structure. This is where we will collect community feedback through comments. [View example]

Project narratives

Our research is generally broken up into projects, concerted efforts to answer specific biological questions, solve particular problems with technology, or sometimes simply to explore a curiosity.

These projects are described on "project narratives." These are intended to serve as a running narratives of progress, almost like a journal or collection of field notes. The project narrative will reflect the current status of the research, where we're heading next, and, sometimes, how you can influence our direction. Project narratives are always evolving, and will not have listed authors or DOIs.

We think they'll be particularly important because we're going to publish much smaller pieces of research than what is traditionally shared in journal articles.

With so many modular publications, or “pubs,” it could become tough to follow the thread. Project narratives will string the individual pubs together into a coherent narrative, summarizing emergent takeaways and laying out planned next steps. Though pubs will work as individual pieces, with all essential context included, the project narratives provide space for more in-depth background.

We’re interested in sharing these narratives, especially the failures and dead-ends, because although these stories are fascinating and informative, they are rarely revealed in published papers. A chronological presentation is more reflective of how scientific work is actually done, and could be useful for students and trainees to study as they set expectations for their own research.

Pub structure and style

Pubs are our standalone “units” of research output.

In **structuring** individual pubs, we’ve tried to strike a balance between improving usability and what scientists expect to see in a research publication. Together, the pub’s title, subtitle/description, and the “Purpose” section provide a high-level summary of key takeaways, features of the pub, and quick links to connected content like the related project narrative, data, or protocols. Section headings will be similar for each “pub type” (e.g. Resource, Observation, Negative Data) so regular readers can get used to their organization, but we will always prioritize making an individual pub as clear as possible. Some sections are similar to those in traditional papers, like “The results” and “The method,” while others, like “What’s next?” feel more like headings you’d see in a news article or blog post.

This informal **style** is intentional. We’ll use scientific terms as necessary to be accurate and to effectively communicate to other experts, but generally aim to strike a more conversational tone to enhance readability. Long sentences and jargon make papers difficult to parse, and we’re hopeful that we can improve the reading experience for all audiences, including technical experts.

The **table of contents** makes pubs navigable, providing a bird’s-eye view of what readers will find in the piece and quick links to jump to the desired information. Right now, users can open the table of contents by clicking the “CONTENTS” button on the right side of each pub.

This is a feature that we'd like to improve upon in subsequent iterations of the site. We've made the table of contents "sticky," so it's always available no matter where within the pub a reader is looking, but we suspect that having it pre-expanded will be much better, especially for those new to the platform. We welcome other suggestions on how to navigate our pubs, especially any examples of sites that do this well.

Non-text elements like figures, code, or other media will appear as close to where they're discussed as possible to minimize scrolling. Whenever possible, we will directly embed content so readers can consider it in context and don't have to juggle multiple open windows. Every reference to a figure also contains a link that allows jumping directly to that figure.

This is another area where we're already imagining future experiments in readability and utility. What if figures always followed you when reading the section in which they're described? What if you could view figures and text in two separate columns? What if individual elements within figures were clickable and could change to display data differently or show more/less information? We'd again appreciate thoughts on features you'd appreciate, lessons learned from previous attempts at improving figures and other media elements, and examples of people or sites making progress in this area.

Finally, pubs are not static — we will revise them iteratively based on feedback. We've decided to start with a **three-stage tag system** to indicate where we stand in the iteration cycle. At the top of each pub, you'll see either "*Feedback requested*" when we've just posted it and are awaiting constructive criticism, "*Revised after community feedback*" when we've updated the pub based on comments, or "*Revised, no longer actively updating*" once the feedback tapers off and any additional comments are better addressed through a new pub.

You can see all release versions of a pub by clicking the rewind/clock icon in the header. When changes more substantial than adjusting formatting or correcting typos are posted, we will include a release note that describes the nature of the changes.

Sharing protocols, data, and code

To make our work as useful to others as possible, we are planning to share as much practical, usable material as we can. This means sharing full protocols, raw data, executable code, and more. We'll strive to always keep our audience and goals in mind when sharing a pub — if we want our readers to be able to replicate our work or apply a method in their own research, we need to include sufficient detail, especially for those whose expertise differs from our own.

When adapting or developing new methods, we aim to post **detailed, step-by-step protocols** on protocols.io. This pub contains an example. We've embedded the protocol directly so it's easy for readers to skim and determine if it's relevant, and they can of course also view and use it in a separate window as well.

Our **data** will be deposited in stable repositories that are findable, accessible, interoperable, and reusable (FAIR) [2]. The goal is for data to meet standards that will maximally facilitate interaction by both humans and machines, further increasing the reach and impact of the open data. In some cases, we may not have yet identified an appropriate repository, or deposition may be in progress. To avoid delays in sharing, we may first share our datasets directly and update links once repository deposition is complete.

We want to share **code** in ways that make it straightforward to run and make outputs reproducible. This goal is more achievable than ever as cloud computing becomes widespread and browsers support embeddable, in-line programming environments. For example, we are taking advantage of tools like Google Colab notebooks or Github gists, which allow users to easily edit and run code on their own or view code directly within our pubs. We'd eventually like to find a reliable way to share executable code within pubs as well. This pub is a great example of this theory in action, and we appreciate any feedback on its utility.

Citation style

Our chosen citation style is minimal, dropping issue/page numbers and journal titles. While non-experts on a specific topic need to know whether experts have

weighed in on scientific research, science is often inappropriately gauged using journal names as a proxy for scientific quality. This is typically based on their presumed level of selectivity or due to inappropriate application of journal-based metrics [3] to individual research articles, as is frequently done with journal impact factors. To reduce the dependence on poor proxies of scientific quality, we have removed journal names from our citations. For simplicity, we have also removed issue and page details, which are less relevant in a digital context, but retained DOI links to make our cited references easily findable. Readers can easily see a citation by hovering over the bracketed citation number (desktop) or tapping it (mobile), and can then click the URL to see the paper.

Other PubPub users can adopt this approach by selecting the “Arcadia Science” style in PubPub’s dropdown list of citation styles. Outside of PubPub, you can access this simplified citation style by downloading the Citation Style Language (CSL) file on our public Github. It will eventually be available in the official CSL repo, pending review, which will make it a style option in many citation management tools.

Authorship, credit, and responsibility

To facilitate transparency and collaboration, we are capturing the substance of individual contributions using a modified version of the contributor roles taxonomy (CRediT). The roles include areas like conceptualization, editing, investigation, methodology, supervision, and more. These roles are indicated next to each alphabetically ordered contributor at the bottom of our pubs. We are extending the taxonomy to include key players not typically included in author lists, such as public commenters who may play an instrumental role in shaping the direction of our work.

The subset of these contributors who are *responsible* for the content of a pub will have signed off on the released versions and will be “byline” contributors. Whether a contributor is considered byline is determined by the set of roles that they played. We are deciding (and plan to share in upcoming communications) which roles convert a contributor to byline by default, so the system can be applied to each new pub with minimal conflict or confusion. Byline contributors will

technically be included as the “authors” in the metadata for the DOI provided through DataCite. These are the names that will appear in citations of our work.

Why make this distinction? What does it mean to be “responsible” for scientific research? When it comes to responsibility and accountability for pubs, we are aligned with the recommendations from the International Committee of Medical Journal Editors (ICMJE), in which byline contributors are accountable for their own work as well as aware/can vouch for the integrity of the contributions of other byline contributors. The byline contributors are similarly the first point of interaction for responding to public feedback and they may engage other contributors as appropriate.

Community feedback as a means of peer review

How will we and others be able to gauge the quality and rigor of our work? Quality control is typically attempted through pre-publication peer review, a process in which manuscripts are sent to journals and experts provide comments that must be addressed or rebutted before an editorial decision is made. While currently considered the gold standard, there’s room for improvement. We’re trying three approaches, outlined below.

First, we want to **solicit feedback from a much broader range of people** than the two or three included in typical journal-led peer review. If a large number of scientists provide feedback, their comments can be more modular, focusing on the specific area of the commenter’s expertise and requiring less time to contribute. The science of today is also highly interdisciplinary, and with community-level feedback, we hope to achieve scrutiny across a greater portion of a publication, from big picture to nitty gritty methods. For example, a cell biologist might comment on the microscopy, a statistician on the study design, and a bioinformatician on the data analysis, with organismal biologists providing appropriate context. Collectively, the feedback can be more relevant, efficient, and useful.

Next, we want to **make all feedback public**. While we may not engage with every comment, the contributions of those commenters will still be visible to all and it will be in the interest of our scientists to attend to critical questions. All readers can then learn from and consider the opinions shared, especially those by experts outside of their domain. Beyond PubPub comments, we will also try to capture public commentary on our work from social media.

Finally, this open **evaluation will happen over time** rather than at a single point, better matching the reality that science is a living, breathing body of work. A rigorous finding is one that stands the test of time, is built upon, and is challenged by orthogonal experimentation. We will solicit feedback, address it, and solicit more feedback. Eventually, we will no longer actively check or update pubs, but comments will always remain open.

Iterating on public review and learning along the way

The scientific literature is growing exponentially [4] and mechanisms for quality control and improvement *must* adapt with it. Crafting sustainable and scalable mechanisms of peer review may be the trickiest piece of the publishing puzzle, and we'll need to test multiple solutions. Here are two major issues we anticipate and initial ideas for how to deal with them:

Will the scientific community participate? Will experts provide sufficient comments to yield confidence in our work without traditional editorial motivators? The heavily managed process in which journal editors with stature directly request reviews cannot be replaced by shouting new science into the void and hoping people will notice. We, too, will need to actively solicit feedback. Now that our initial content is available, we're beginning to explore ideas to cultivate engagement with our work, and hope to relay our learnings in future pubs. We welcome ideas too!

Without the pre-existing audience of a journal, how will we attract interest at all? We hope that, in addition to actively soliciting reviewers, doing the kind of work that matters to people will help us build an invested community. We will have to produce work that can easily be used and improved by others. We must demonstrate true impact over time, and welcome feedback from all who might benefit from our contributions. We must also be responsive to the community's

bigger-picture direction, being willing to re-chart our scientific path to better serve the greater need.

Our own contribution to public review

Why should anyone donate their time to give public feedback to a company? We aim to be good community members in return.

First, we aspire to generate the kind of data, insights, and technologies that can dramatically advance what scientists in the community are able to do in their own work. Feedback will just make the quality of our contributions and responsiveness to the needs of the community that much better.

Second, we are asking our scientists to actively participate in commenting on preprints and other public research products at a level commensurate with their own research output. Arcadia will appear as a reviewer group on Sciety and our comments on preprints will be displayed there and on bioRxiv. We hope our comments help the authors of those products improve their own work such that we can contribute to the collective advancement of knowledge and earn their valuable feedback.

Standing on the shoulders of the open science community

While we have the unique opportunity to run agile publishing experiments through a growing research organization, this first version of this model was inspired by the work of many prior and forthcoming open science efforts. These include the preprint servers arXiv, biorRxiv, OSFpreprints, and so many others. Our efforts are also inspired by the F1000 research platform for post-publication review, the "publish, then review" efforts at eLife based on a publish-review-curate model, crowd review efforts at ASAPbio.org, community preprint review efforts such as PreReview and Peer Community In, and activities of many other organizations represented at Incentivizing Collaborative Open Research (ICOR), who are innovating on open and reproducible efforts throughout every stage of the research cycle.

Acknowledgements

We are grateful to everyone who has contributed to this first iteration of our publishing site. Thank you to Mert Celebi for weighing in on countless technical questions and developing solutions to key challenges. Thank you to Seemay Chou for all your thoughtful insights and for making the time to help pull everything together in the past few weeks. Thank you to Gabe Stein and everyone on the PubPub team for being great thought partners and making so many site adjustments for us. Thank you to the whole And—Now team for creating and implementing fantastic designs incredibly quickly. Finally, thanks to all the scientists at Arcadia for your amazing ideas and patience over the past few months! Special shout-out to our publishing pioneers, Peter Thuy-Boun and Jase Gehring, for being brave enough to be the first of our scientists to release their work in this new way.

Competing interests

P.A. is the president of ASAPbio and serves on the board of directors at eLife, groups that are also experimenting with scientific publishing.

Contributors (A–Z)

- **Prachee Avasthi**: Conceptualization, Editing, Supervision, Writing
- **Megan L. Hochstrasser**: Conceptualization, Editing, Writing

References

1. Vale RD. (2015). Accelerating scientific publication in biology. https://doi.org/10.1073/pnas.1511912112
2. https://data.europa.eu/doi/10.2777/1524
3. Larivière V, Kiermer V, MacCallum CJ, McNutt M, Patterson M, Pulverer B, Swaminathan S, Taylor S, Curry S. (2016). A simple proposal for the publication of journal citation distributions. https://doi.org/10.1101/062109
4. Bornmann L, Haunschild R, Mutz R. (2021). Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. https://doi.org/10.1057/s41599-021-00903-w