EdTech Hub

Clear evidence, better decisions, more learning.

Disseminating the Evidence and Outputs Generated by Your Programme

Three options for setting up an evidence library

UK International Development — Partnership | Progress | Prosperity

THE WORLD BANK

unicef for every child

About this document

Recommended citation	Haßler, B., Mansour, H., Friese, L., & Longley, S. (2024). *Disseminating the Evidence and Outputs Generated by Your Programme: Three options for setting up an evidence library* (Helpdesk Response No. 178). EdTech Hub. https://doi.org/10.53832/edtechhub.1001. Available from https://docs.edtechhub.org/lib/PWN42VDQ. Available under Creative Commons Attribution 4.0.
Licence	Creative Commons Attribution 4.0 International https://creativecommons.org/licenses/by/4.0/ You — dear readers — are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material) for any purpose, even commercially. You must give appropriate credit, provide a link to the licence, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
Reviewers	Jamie Proctor, Arjun Upadhyay

About EdTech Hub

EdTech Hub is a global research partnership. Our goal is to empower people by giving them the evidence they need to make decisions about technology in education. Our evidence library is a repository of our latest research, findings and wider literature on EdTech. As a global partnership, we seek to make our evidence available and accessible to those who are looking for EdTech solutions worldwide.

EdTech Hub is supported by UKAid, Bill & Melinda Gates Foundation, World Bank, and UNICEF. The views in this document do not necessarily reflect the views of these organisations.

To find out more about us, go to edtechhub.org/. Our evidence library can be found at docs.edtechhub.org/lib/.

Contents

1. Getting your research into circulation: Setting up an evidence library

This document is designed as a guide for building and implementing a content management system (CMS), specifically an evidence library, prioritising the needs of research programmes generating evidence for other researchers and policymakers and for public readership and global dissemination. The primary focus is on securely building a system to meet research discovery requirements, such as search engine optimisation (SEO) and visibility on Google Scholar.

A robust content management system like an evidence library goes beyond simply meeting web standards. It is about enhancing the *research discoverability* of your work, gauging your impact and reach, increasing the accessibility of your evidence, and fostering a more collaborative environment. This comprehensive approach is vital for successfully sharing evidence with the public and the global community.

The document is primarily aimed at project or communications leads. We assume no specific technical expertise. It seeks to help you decide how to deploy an evidence library and whether to draw on internal or external resources (such as web developers). We expect the document to also be useful for product owners and web developers; however, we expect that for such users, the document will cover a lot of familiar ground.

The tools we advocate for are all open source, and in this document, we will describe three options:

1. You can host the software yourself.

2. You can hire a developer to host the software for you.

3. You can use the turnkey solution available (https://opendeved.net/hostedevidencelibrary/).

1.1. Problem statement

Most large-scale development programmes set up their own websites. Typically, you would use a readily available web content management system such as WordPress to host your website. It would be unusual to build a website without such tools. Discovery of your website in search engines, such as Google, relies on search engine optimisation—which is an

area well understood by web developers; most content management systems have excellent tools and plugins for search engine optimisation.

Publications — your journal articles, preprints, reports, research briefs, conference presentations — and their discovery follow similar rules. They need to be optimised for discovery in publication search engines using tools such as Google Scholar. However, in contrast with general websites, the 'content management systems' and 'search engine optimisation' for publications are usually poorly understood by web developers and remain the purview of specialists working for journals or in the field of bibliometrics.

Consequently, many publications from major research programmes remain undiscovered, even when significant investments are being made. Moreover, even if publications are discovered, they are not cited uniformly. This makes it difficult to track the publication back to the original research programme, making it difficult for the research programme to measure impact.

In summary, you may struggle with the following issues:

- My publications don't get discovered by search engines.

- My publications aren't citable.

- I can't track the use of my publications.

1.2. Initial recommendations

Here are two important initial recommendations:

1. Open-source systems are available for hosting both web content and publications. Under no circumstances should you develop a system 'from scratch'.

2. Deploying a standards-based website — including publications — can initially cost you between thousands and tens of thousands of GBP, with significant recurring costs. Sometimes, budgets in the hundreds of thousands are allocated to web content hosting; such high costs are unnecessary and avoidable.

2. What is an evidence library?

Several actors in the development space, including EdTech Hub, Open Development & Education, and ESSA Africa (Education Sub-Saharan Africa), have pioneered the concept of 'evidence libraries'. Among other goals, the purpose of an 'evidence library' is to:

1. Include and make available your own programme's diverse range of outputs, listed in a standards-based way that enhances discovery.

2. Display and provide access to outputs in your wider content management context. An evidence library can offer the user a seamless experience between your web content and your publications.

3. Make it easy for your staff and other researchers to cite and build on your publications.

4. Optionally, showcase your outputs in the context of wider research and grey literature: As well as curating outputs from your programme, you can curate other relevant research, making this easily accessible to your audience. This includes showcasing publications you cite, as well as tracking and showcasing who cites you.

In developing software tools for evidence libraries, we consider other efforts, such as content management for publications and preprints. The specific components that we use to build 'evidence libraries' meet the above four criteria.

Figure 1. *Snapshot of the EdTech Hub evidence library* (https://docs.edtechhub.org).[1]

However, if your concerns are different, particularly in the field of natural sciences, you may wish to consider alternative tools. It is worth noting that the goal of an evidence library differs from that of a preprint server; a preprint server typically offers a mechanism for publishing preprints (from multiple programmes) on route to journal publication. The primary purpose of an evidence library is to aggregate your programme outputs, including any outputs that may have been published on preprint servers. If you wish to set up a preprint server, web-based research management tools are available. The Open Science Foundation offers an integrated research management and preprint tool (https://osf.io/)[2] that is worth considering. Preprint servers are common in some areas, particularly natural sciences, but — regrettably — less common in education.

---

[1] See https://docs.edtechhub.org. Retrieved 27 August 2024.
[2] See https://osf.io/. Retrieved 27 August 2024.

3. The central components of our evidence library: Zotero and Kerko

Research in education, social science, and humanities typically makes extensive use of publication management and literature review. Among others, a common open-source tool for publication management and literature review is Zotero.[3] An important feature of Zotero is that it offers collaborative libraries; such shared libraries can allow researchers in your programme to share relevant publications. Particularly, when working with international teams, using collaborative Zotero libraries can keep your discussions grounded in evidence.

For our evidence library system, Zotero is the central information component. Zotero is a highly sophisticated and flexible system for evidence acquisition and dissemination that is tailored for managing publications and information. This has both advantages and disadvantages. Advantages include:

- Zotero is designed for the curation of evidence, so it will become a central piece of your infrastructure, organically organising your research outputs, and showcasing these internally as well as externally.

- Zotero is also a reference manager, so you and your team can easily cite research in publications.

- The use of Zotero will promote research literacy within your teams, making your programme more evidence-based, and, therefore, more efficient and effective.

Figure 2. *Data flow from the Zotero application to the Zotero cloud datastore to the web application.*

Overall, it is extremely unlikely that your programme will outgrow the facilities provided by Zotero. However, because Zotero is flexible and

---

[3] See https://zotero.org. Retrieved 27 August 2024.

sophisticated, there will be a learning curve, particularly for staff unfamiliar with education research processes and reference management.[4]

Zotero stores information in the cloud and offers an API (Application Programming Interface). This API uses the Kerko tool[5] to generate standards-compliant web listings.[6]

---

[4] Extensive documentation and tutorials for using Zotero are available.

[5] See https://github.com/whiskyechobravo/kerko. Retrieved 27 August 2024.

[6] Collaborative Zotero libraries can also be accessed online through a web interface, and permissions can be set so that such libraries can be accessed publicly as read-only. However, this web interface is similar to the application interface, and is intended for other researchers, rather than public dissemination.

4. Options for deploying an evidence library

This section contains significant details to support you in deciding how to host an evidence library. The section is aimed at programme leads or communications managers who need to make decisions about deployment options without insight into all the technical details involved. As a product owner or web developer, you will find this section less useful; however, you may wish to look at the various links to the software with further instructions.

Because the system is open source, we note that the initial costs are very low. The Zotero application is open source and free to use; basic Zotero data storage is free and very low cost, even for extensive storage requirements. Kerko is open-source and can be deployed on any server. However, deploying the system will incur some costs. For example, you will need some infrastructure, and will likely wish to tailor the Kerko software, e.g., to meet your requirements.

For hosting the open-source Kerko software, you have the same options as for hosting any open-source software. Therefore, while we have tailored this section to our use case, many of the instructions apply to any open-source software.

You have three options, broadly speaking. We have listed them together with questions to help you decide.

1. **You can host the Kerko software yourself.**
   – Ask yourself: *Does your project already self-host some software? Do you have staff familiar with the requirements for self-hosting?*

   – If the answer to these questions is yes, consider self-hosting. Otherwise, consider the other options.

2. **You can hire a developer to host Kerko for you.**
   – Ask yourself: *Does your project already employ and manage external developers?*

   – If the answer to this is yes, you could consider hiring a new developer with the requisite skills. If your project has not engaged with developers yet, we suggest considering the final option.

3. Use the turnkey solution available.

– If you are unfamiliar with the first two approaches, a turnkey solution might be the most appropriate. Such turnkey solutions are not available for all open-source projects; however, for an evidence library, a turnkey solution is available from the original developers of the EdTech Hub evidence library.

– Specifically, Open Development & Education,[7] a consortium member of EdTech Hub, offers a turnkey solution for tailoring and deploying evidence libraries (further details: https://opendeved.net/hostedevidencelibrary).[8] The approach uses Zotero and Kerko, as described in this document. In line with the Principles for Digital Development on open source, note that there is no lock-in: All data and infrastructure can be migrated to a self-hosted solution anytime. Open Development & Education also provides training for all aspects described in this document.

These three options for deploying Kerko are analogous to deploying Tangerine,[9] or, indeed, many other open-source approaches. The next sections provide information on the three approaches and further details on selecting the approach that suits your needs.

---

[7] See https://opendeved.net/. Retrieved 27 August 2024.
[8] Retrieved 27 August 2024.
[9] Host Tangerine yourself: https://github.com/Tangerine-Community/Tangerine; access Tangerine as a cloud-hosted service from RTI: https://www.tangerinecentral.org.

Figure 3. *Decision tree: Do I need an evidence library?*
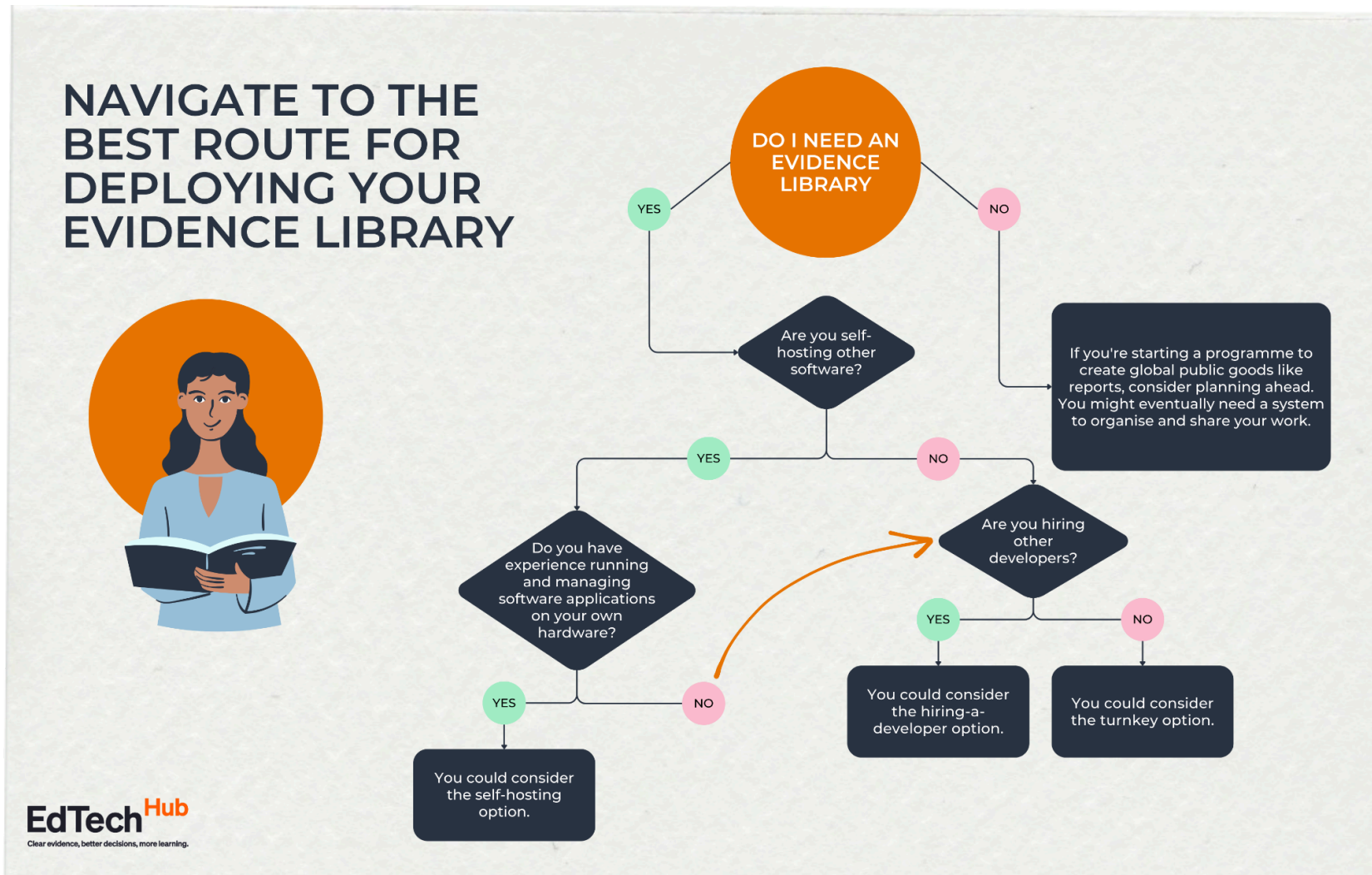
4.1. Considerations for self-hosting

To get the full system working step by step, you have to undertake the following:

- Sign up for a Zotero account at https://zotero.org[10] (free)

- Set up a Zotero group library for your programme (and enable shared data storage)

- Host and deploy Kerko with API access to your Zotero group library.

The first two steps are relatively straightforward. However, hosting the Kerko software requires significant technical expertise. If you have the necessary technical expertise on your team, you can host the Kerko software yourself. The Kerko GitHub repository provides full instructions on how to run the Kerko software: https://github.com/whiskyechobravo/kerko.[11] You will need to acquire a server, e.g., based in a data centre that you have access to, or use a cloud-based service.

The remainder of this section covers the details of what self-hosting involves. This can help you decide whether self-hosting is suitable for you. As noted above, if this is the first time you are self-hosting software, this may not be the right approach for you. However, if you regularly self-host software for your programme, this option could work for you.

Key terms

- **Open-source software**: This is a type of software where the source code — the part of the software that tells it how to work — is available for anyone to see, modify, and distribute. This is different from proprietary software like Microsoft Office or Adobe Photoshop, where the source code is kept secret by the company that created it. Open-source software is often developed collaboratively by a community of users and developers.

- **Self-hosting**: This term refers to running software on your own computer or server, rather than using a version that's hosted on someone else's server. For example, if you use Google Docs, you're using software hosted on Google's servers. Self-hosting would be like setting up your own version of a document editor on your own computer or a server you control.

---

[10] Retrieved 27 August 2024.
[11] Retrieved 27 August 2024.

Putting these two concepts together, self-hosting open-source software means that you use software with an openly available code and run it on your own computer or server. This gives you direct control over the software without involving a third party. However, self-hosting also means you are responsible for setting up the software, keeping it running, and dealing with any issues that arise. This can be challenging for non-technical users or for projects that do not have appropriately qualified staff.

Self-hosting software in detail

Some considerations for self-hosting software:

- **Prepare your hardware**

    - For Kerko to run, you will need a cloud-based server or a server in your own data centre if you have one.

- **Install an operating system**

    - Install a server operating system: e.g., Ubuntu Server or Debian.

    - Secure your operating system with updates, firewalls, and other security best practices.

- **Install and configure the software**

    - Follow the installation instructions provided by the software. This involves downloading the software, setting up databases, and configuring settings. The Kerko GitHub repository provides full instructions on how to run the Kerko software: https://github.com/whiskyechobravo/kerko.

    - You will need basic command-line operations because the software does not have a graphical installer.

- **Maintenance, resource management, data management, and backup**

    - Once the software is running, you will need to monitor the server. Some Kerko processes can be resource-intensive. You will also need to devise a backup strategy.

    - You must regularly update the software and operating system for security and feature enhancements.

    - Monitor the performance and health of your server.

Overall, consider:

- **Technical skill**: Do you have the skills required for self-hosting?

- **Time and effort**: Do you have the time available to set up and maintain a self-hosted server?

- **Security**: Can you mitigate any security issues that arise?

Self-hosting open-source software can be a rewarding project. However, it is important to weigh the benefits against the time, effort, and technical skill required.

4.2. Considerations for hiring a developer

You can hire a developer to host Kerko for you. Many sites are available where you can hire developers (such as Upwork); few developers will be directly familiar with Kerko; you would need to look for a developer familiar with Python.

Deploying open-source software on a server can be a complex task, especially if you do not have in-house technical expertise. Hiring a developer can be a great solution. This section lists essential considerations to bear in mind to ensure you find the right person for the job.

We do not recommend this option if you have never hired software engineers or software developers for your programme. However, if you regularly hire developers for your programme, this option could work for you. Here is a list of considerations.

Technical expertise and experience

There are several skills and kinds of technical expertise you should look for.

- **Relevant skills**: Look for a developer with experience in the specific technologies and programming languages used by the open-source software you want to deploy: Kerko, Zotero, python, nodejs. While there is a vibrant community of developers who are familiar with Kerko and Zotero, it can be hard to find developers who are available for hire.

- **Server management**: Ensure the developers have experience with server management, including setup, configuration, security, and maintenance. This requirement will be fulfilled relatively easily.

- **Portfolio / previous work**: Ask for examples of similar projects the developers have worked on. This can give you an idea of their capability and experience. As with the first bullet point in this list, this may be difficult.

- **Understanding of open-source software**: The developer you hire should have an understanding of open-source software. This requirement will be met relatively easily.

Further, many general considerations for hiring contractors also apply to hiring a software developer. For example, when hiring a software developer for server deployment, the person should have clear communication skills and should be able to simplify technical concepts. They should be adept at managing project timelines and deadlines, and possess problem-solving skills for any deployment challenges. Knowledge of security best practices and data protection, especially for sensitive data handling, is essential. Understanding a developer's cost structure, whether hourly, fixed, or milestone-based, and clarifying payment terms is important to align with your budget. Post-deployment support, including maintenance, updates, and troubleshooting, as well as comprehensive documentation of the setup, are essential for ongoing operations.

As with any subcontract, you need a clear, written contract outlining the scope of work, deliverables, timelines, payment terms, confidentiality, and any other important terms. We note that you should ensure that the resulting code becomes your intellectual property and should be released under an open licence.

Considering these factors can increase your chances of hiring a developer who can successfully deploy the open-source evidence library software on your server.

4.3. Considerations for using the Kerko turnkey solution

Self-hosting open-source software and using turnkey solutions are two different approaches to deploying and using software, each with its set of advantages and disadvantages. Broadly speaking, if you have no experience with self-hosting software in your programme and you do not regularly hire software developers, it may be better to use a turnkey solution: This is likely to save you time and effort and get you up and running much faster. A turnkey solution is available from Open

Development & Education; see
https://opendeved.net/hostedevidencelibrary/.

Advantages of self-hosting open-source software

- **Control and customisation**: If you have appropriately qualified staff on your team, you have complete control over the software and can customise it to meet your specific needs; this includes the ability to modify the code, integrate with other systems, and configure settings to your preference.

- **Privacy and security**: Self-hosting can offer enhanced privacy since your data is stored on your own servers. You are also in control of your security measures, which can be a significant advantage if you have the expertise to implement strong security protocols.

- **Cost**: If your organisation relies on many turnkey solutions based on open-source products, you might find that these can be more expensive over time, especially with subscription-based models.

Advantages of the evidence library turnkey solution

- **No technical expertise required**: As noted above, setting up and maintaining a self-hosted environment requires a certain level of technical skill. This can be a significant barrier for those without software development experience. The evidence library turnkey solution is designed for ease of use, with minimal setup requirements.

- **Professional support**: Unlike commercial software, open-source projects may lack dedicated professional support. Assistance is typically community-driven and might not always be readily available or timely. The evidence library solution comes with professional support and customer service, providing help for troubleshooting and technical difficulties.

- **Quick deployment**: The turnkey solution can be deployed quickly without significant effort from your side.

- **Ongoing maintenance and updates are provided**: You do not need to focus on updates and maintenance, reducing the workload on your programme.

- **Controlling your costs**: While the open-source software is free, the software cost is only part of the total cost. You need to consider the costs of setting up the software and the hardware, and the cost of

maintenance. Overall, those costs could be higher than the cost of using a turnkey solution.

We note that while there is always the risk of service discontinuation, there is no vendor lock-in: All software is open source, and all the data is yours and can be migrated at any time. Furthermore, you can switch between the above-mentioned three options at any time: For example, to speed up setup, you could start with the turnkey solution; if and when desired, you could switch to a self-hosted solution at a later stage.

4.4. Conclusion

Overall, you will benefit from whichever of the three approaches you choose, including cost savings. In all three cases, you are building on existing systems, which will speed up the deployment of your evidence library.

You will benefit from discoverability and standards-based listings for a lower cost than developing a full system from scratch. Rather than focusing on new developments, your entire effort can be focused on tailoring the evidence library to your needs. Where well-maintained open-source solutions exist, as with Kerko and Tangerine, it is simply not cost-effective to develop full solutions from scratch.

5. Additional features available

As mentioned, the key components for setting up an evidence library are Zotero and Kerko, and you can deploy your system based on those components alone. However, depending on your use case, additional features are available.

5.1. Automated DOI allocation (My Education Evidence)

The My Education Evidence service (https://my.educationevidence.io)[12] is a web interface enabling several important features. My Education Evidence is part of the turnkey solution for tailoring and deploying evidence libraries. While the setup is more complex for various reasons, My Education Evidence is also open-source software and could be self-hosted by your organisation.

Figure 4. *Data flow from the My Education Evidence backend to Zotero datastore.*
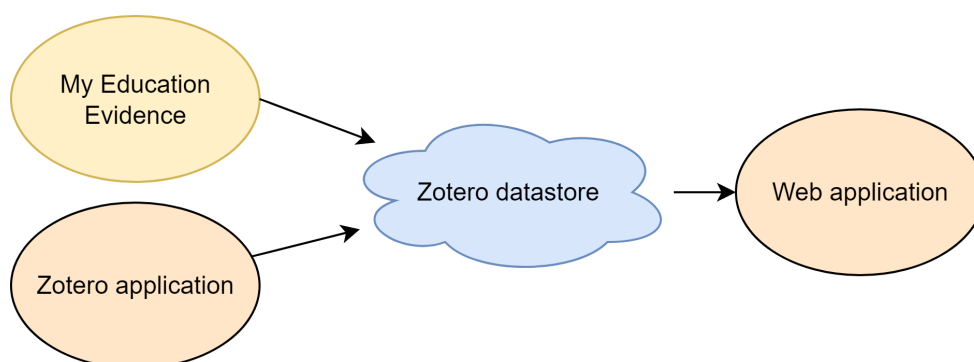
My Education Evidence offers a user-friendly web interface for registering organisational outputs, as well as tagging them. The web interface would typically be available to your entire organisation. It assists with capturing all outputs generated by your programme. For example, systematic capture of EdTech Hub outputs has drawn attention to the fact that an output is registered every few days, totalling some 100 outputs per year. With My Education Evidence, these outputs are automatically tracked internally.

---

[12] Retrieved 27 August 2024.

Figure 5. *Screenshot of example record listings in My Education Evidence.*
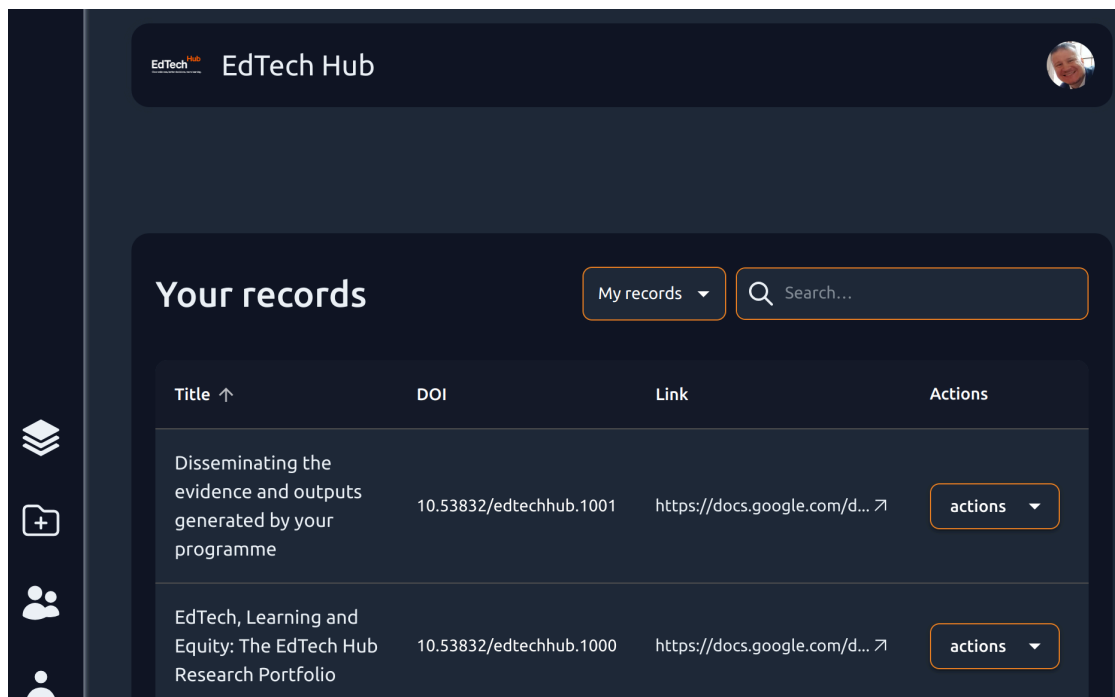
The service also offers Digital Object Identifier (DOI) allocation with Crossref. In this way, all of your outputs (including reports, preprints, research briefs, conference presentations, etc.) are registered in Crossref.[13] Crossref is the universal register for research publications and is used by all high-profile journals. By using Crossref, your outputs become part of the Crossref ecosystem. This significantly enhances the discoverability of your outputs: They can be cited more easily and are more likely to be picked up by search engines (such as Google Scholar). Because the DOI gives you a short, unique label, you can also use the DOI to find out where your publications have been cited.

Additionally, as Crossref registration requires, all outputs are archived in a public repository for long-term archiving (Zenodo). Such archiving is routinely undertaken by universities; however, many large-scale donor-funded projects do not adequately archive programme outputs, leading to their loss after the programme's closure.

---

[13] Crossref is a non-profit open digital infrastructure organisation for the global scholarly research community. It uniquely and persistently records and connects knowledge through open metadata and identifiers for all research objects such as grants and articles. See also https://en.wikipedia.org/wiki/Crossref, retrieved 27 August 2024.

Figure 6. *Record creation in My Education Evidence. The interface can be tailored. For EdTech Hub, certain categories are assigned at the time of registration.*

Figure 7. *The tagging interface — the image shows some of the categories available for EdTech Hub outputs. The selection of tags/taxonomy is specific to each programme and can be adapted as needed, as noted below in* Section 6*.*
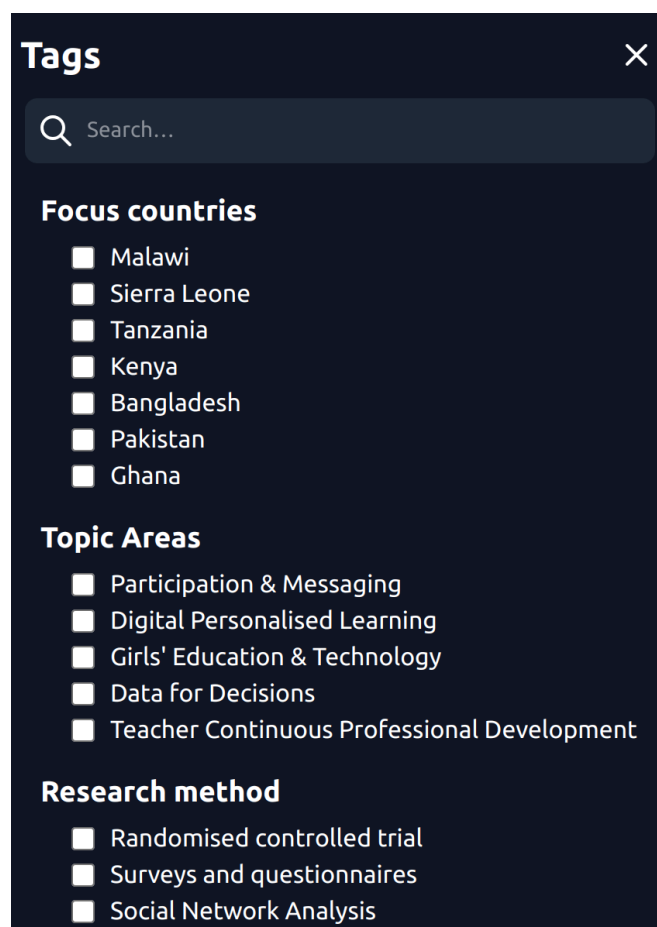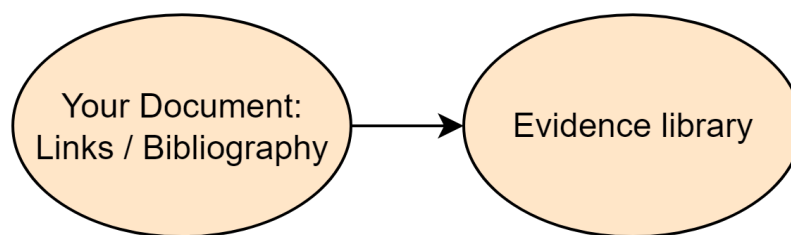

5.2. Zotero plugins

A key feature of Zotero is referencing and citation. The researchers in your programme may already be familiar with using reference managers like Zotero. As part of the evidence ecosystem described in this document, several plugins that help promote your evidence are available.

The way these plugins function is very similar to the usual plugins for citation and bibliography generation. However, our plugins create (trackable) links between your documents and your evidence library.

Figure 8. *Function of the Zotero plugin.*

Our current plugins include a Zotero plugin; the plugin allows you to copy references to your clipboard, ready for pasting into other documents, such as Google Docs. A plugin is also available for Google Docs and Google Slides, facilitating linking documents to your evidence library and generating bibliographies.

For example, if a publication is cited in the present document (e.g., ⇧Haßler et al., 2021), the publication is listed on the evidence library record for the present document (https://docs.edtechhub.org/lib/PWN42VDQ) under '**Cites**'. The Google Docs plugins also make it possible to automatically generate bibliographies or reference lists from the citations embedded in a document, as shown below.

Bibliography

This bibliography is available digitally in our evidence library at https://docs.edtechhub.org/lib/PWN42VDQ

Haßler, B., El-Serafy, Y., Khalayleh, A., & McBurnie, C. (2021n). *Systemic Mixed-Methods Research—a conceptual framework for scalable EdTech research* (Working Paper 9). EdTech Hub. https://doi.org/10.53832/edtechhub.0001. Available from https://docs.edtechhub.org/lib/TD2ASC4G. Available under Creative Commons Attribution 4.0 International. (details)

WordPress

If you use WordPress to host your website, a plugin for connecting your WordPress site to your evidence library is also available, allowing you to create and embed bespoke bibliographic lists in WordPress pages.

6. Considerations for managing your content

It is important to note that the above description focuses on technologies for managing publications. While these technologies give you the tools to manage your programme's publications, you still need to manage your content and develop strategies accordingly.

For example, you may need to devise a set of categories that make sense to your users. Some programmes organise their content using categories related to programme management. While this may make sense to programme staff, the rationale may not be clear to external users. The systems outlined in this document allow you to have both 'internal categories' (useful for programme management, visible to programme staff only) and 'external categories' (that appear on your evidence library, visible to internal and external users).

EdTech Hub uses various categories; for example, there is a category for types of learners:

- Girls and gender

- Learners with disabilities and special educational needs

- Low-level foundational skills

- Minority groups

- Out-of-school populations

- Refugees, migrants, and displaced persons

Among many other categories, EdTech Hub also uses 'topic area' and 'focus countries', which align with the overall goals/remit of the programme (i.e., 'data for decisions' etc., as core research topics, while Sierra Leone, Pakistan etc., are focus countries).

- Data for Decisions

- Digital Personalised Learning

- Girls' Education & Technology

- Participation & Messaging

- Teacher Continuous Professional Development

The list of focus countries includes:

- Bangladesh

- Ghana

- Kenya

- Malawi

- Pakistan

- Sierra Leone

- Tanzania

Note that within the systems outlined here — Zotero, Kerko, My Education Evidence — there is a lot of flexibility regarding the content organisation. For example, changing, adding, or merging categories is straightforward.