



Lighthouse for the Lost – Applying Discovery Tools to Lower the Usability Barrier for Research4Life Institutions

Kimberly J. Parker
World Health Organization
Geneva, Switzerland

Stephen A. Gilewski
Serials Solutions
Seattle, WA, USA

Session:

149 – Evolving serials – managing, discovering and supplying highly innovative and dynamic content – Serials and Other Continuing Resources

Abstract:

Research4Life is an initiative that brings vital scientific, technical, and medical knowledge to researchers and others in the world's least developed countries. Since the HINARI programme first launched in 2001 the concept has been replicated in three other topical areas and the Research4Life partnership now includes four United Nations (UN) entities, two United States academic libraries, content from over 160 publisher partners, and technical expertise and services from companies such as Microsoft, Swets, and Serials Solutions.

Serials Solutions joined the partnership in 2011 and the door was opened to a range of new possibilities. To date, the programmes have depended on traditional A&I (abstracting and indexing) tools such as PubMed for identifying relevant content. However, much of the content accessible by Research4Life is not indexed by these traditional tools. In addition, there was the challenge of usability. The majority of those in low- and-middle-income countries (LMICs) using the Research4Life programmes are casual, or novice users. Navigating a number of distinct systems in order to discover and access content thus has proved difficult for them. With Summon, (Serials Solutions' web scale discovery tool) the Research4Life users will now be able to conduct advanced research from one single starting point.

The challenge for Research4Life and Serials Solutions has been how to set up an enormous consortium with 106 countries, areas, and territories accessing the

content. The professional librarians on staff at Serials Solutions jumped at the opportunity to contribute, volunteering their personal time to enable this mammoth setup. All in the spirit of advancing research around the globe, the plan developed was for each volunteer to “adopt a country.” This has helped to further focus efforts and foster more personal connections to the project.

Challenges remain and the opportunities for LMIC users are only just beginning to be realized, but the journey has already been an exciting one.

Background

Research4Life began with the HINARI Access to Research in Health programme that was conceived in early 2001 to offer developing country institutions free or nominal-cost online access to the world's biomedical literature. Ten years later, HINARI is now replicated in agriculture (AGORA), the environment (OARE), and applied sciences for innovation (ARDI). The partnership has grown to include four United Nations entities, Yale and Cornell universities, more than 190 publishers, the International Association of STM Publishers, Serials Solutions, Microsoft Corporation, Swets Information Services, and innumerable individuals and institutions who have become champions and ambassadors of the concept.

Each of the four Research4Life portals enables users from subscribing institutions to search publishers' resources, view abstracts of publications, and download the complete texts of journals and books.

HINARI makes available over 8,000 journals and 7,000 books to more than 5,000 organizations, including universities, teaching hospitals, research institutions, government offices, and national non-governmental organizations (NGOs) in 106 countries, areas and territories. AGORA provides more than 2,100 journals and 900 books to more than 1,900 institutions, and OARE includes more than 4,500 journals and 3,500 books to more than 2,400 institutions. ARDI is the most recent programme in Research4Life, having joined the partnership in 2011.

Serials Solutions, a ProQuest business, is a global leader in essential discovery and e-resource management solutions, powered by a comprehensive knowledgebase. Founded in Seattle, Washington, USA in 2000, Serials Solutions was created by a librarian for librarians. The company now has over 230 employees and serves libraries and other institutions in over 77 countries across the globe. Serials Solutions' flagship product is the web-scale discovery service, Summon.

For Serials Solutions, the idea for a Research4Life partnership began with a visit to the World Health Organization (WHO). WHO was exploring the potential of discovery services and upon learning about Summon, noted how useful it would be for the HINARI programme (not to mention the other programmes in Research4Life).

There was immediate interest in a collaborative project from both sides. After all, the goals of the Research4Life partnership and Summon are essentially the same – to help users search, discover, and access content. Further, since the focus of Research4Life is in low- and middle-income countries, such a project would enable Summon to reach users in locations where it would otherwise be inaccessible.

Serials Solutions agreed to donate Summon sites to the partnership, with each country eventually getting their own specific instance. Even starting with only the HINARI programme, this would require a good amount of work. Due to the nature of the project the effort would need to come from volunteers.

The Research4Life partnership was presented during a Serials Solutions' "Librarians who Lunch" meeting, a group developed for librarians within the company to coordinate events and discussions related to the field. Predictably, there was overwhelming interest in the project. It offered a volunteer opportunity to help researchers in developing nations discover content. This touches on many areas of interest for Serials Solutions' librarians. There are elements related to cataloguing and metadata, as well as information access and retrieval. Further, it is a very real way for the librarians to use their expertise for a greater good. The aim is to help users discover and access content, which really is a fundamental crux of librarianship.

Serials Solutions as a company is librarian-focused. Many involved in the organization have graduate library degrees and a deep interest in the field. In this context, the company's mission of serving libraries takes on a more profound meaning as jobs are aligned with passions. With the Research4Life Summon project Serials Solutions volunteers were given the opportunity to "adopt" a member country, and, in essence, be their librarians.

Existing Search Experience in Research4Life

Very soon after HINARI was launched, the US National Library of Medicine (NLM) enabled a tailored version of PubMed that clearly labelled content that was available in HINARI. Users were able to search for relevant articles of interest in the journals covered by PubMed and then assess whether the full text may be available to HINARI. In order to reach the full text, users would have to return to the HINARI user portal, browse through the list of available journals, and then connect through the journal's table of contents. Later, this functionality was supplemented by PubMed's LinkOut option, which allowed users to go directly to the full text of articles.

Similarly, AGORA began with a specialized version of CAB Abstracts, and OARE began with Ebsco's Environment Index and Proquest's Environmental Science and Pollution Management databases. These were later supplemented with additional abstracting and indexing tools, including Thomson's Current Contents, Elsevier's Scopus and many more. These provided enhanced options for discovery, but also added complexity for casual users.

Each of these solutions helped to improve the research process, but still, a contextual explanation and training was required to ensure the most effective information discovery. The casual user was able to discover some content through simple searches, but rarely did they find the "most relevant" or "most complete" content. To facilitate this, the Research4Life programmes created training materials which they continue to update and expand, as well as offering in-person and remote training. For example, HINARI alone has six training modules on PubMed searching, including the complexities and advantages of searching via MeSH (Medical Subject Headings).

Regardless of how extensive are the programmes' efforts, no training can possibly reach all individuals affiliated with the involved organizations (over 6000 institutions). Therefore, many users will remain casual searchers. Therefore, the better the discovery options are for those users, the more the Research4Life provided content will be used for research, policy-making and curriculum building, and in practical settings.

Aside from the difficulties of ensuring the Research4Life user community is able to use the abstracting and indexing tools to their fullest potential, the programmes have a second challenge of discovery. Each programme has a wealth of resources provided by the partner publishers available to the user communities. Many of these resources are not indexed by one database or another. This means that much of the content is not discoverable through the databases, and may only be found by chance or through citations from other articles that may reference the materials.

Summon Approach to Discovery and LMICs

The Summon approach is to achieve for libraries what Google has done for the general public. The goal is to provide a single search box that allows users easily to discover all of an institution's holdings. This can include a library's print and media holdings, in addition to electronic content of all formats (especially journals and ebooks). With Summon, records from library catalogues and institutional repositories are all brought together along with provider and publisher content in a single unified index.

Summon fits perfectly with the needs of Research4Life. Whereas users previously had to learn several disparate systems to discover content, they will now be able to search and discover content from one single starting point. This will reduce the need for extensive user training, as the primary search mechanism will be a familiar and user-friendly interface. The interface can also be changed as required into 30 different languages, with help documentation to match. So, with little or no training, users will be able to search, discover, and access relevant content.

Finding the most relevant articles is a major challenge for the Research4Life users. Previously, a user may have needed to search any number of provider sites sequentially. Each site may have its own index, subject headings, and controlled vocabulary, so search strategies on one site may not work on another. Along with the single search box and unified index, Summon offers a finely tuned relevancy ranking that retrieves appropriate results whether the search is very basic or more complex. This is crucial for the novice user who may not be familiar with the finer points of traditional database searching, such as Boolean operators or controlled vocabularies.

Though one of the main benefits of Summon is its simplicity, it is by no means useful only for the novice researcher. There are many advanced searching options with faceting and sorting available to further refine searches. Users can sort by subject, author, and date. They can limit search results by content type, full text status, and even whether or not they come from scholarly, or peer-reviewed publications. All these options are clearly marked and generally self-explanatory. It is even possible to create discipline-specific scoped search boxes as widgets, which can be added to outside websites.

The scope of the Summon index is huge. At the time of writing, it is approaching nearly one billion records. This can be overwhelming of course, but by default, users will only search the resources they have access to. This is perhaps one of the biggest values of Summon for the Research4Life programmes. By having a profile set up for each country independently there is no guesswork in determining which articles or journals the institutions in that country have permissions / access to. Knowing that everything being searched is accessible in full text online offers great efficiency for the researcher. The time spent researching an article that one will not be able to access is eliminated. Users will no longer have to access multiple sites in order to determine access. Along with this targeted searching though, users still have the option to search beyond their holdings to discover all the content that is in the Summon index.

Summon now offers direct linking to many resources. While the number of resources that are accessible via direct linking is constantly growing, some resources still must be accessed via OpenURL (a context-specific link resolution standard). So, in combination with the Summon site, each country is provided use of 360 Link, Serials Solutions' link resolver. This will further enhance the research experience since after searching and discovering relevant resources, one more click will take

them directly to the full text content. In the end, users will search, discover, and access resources in one seamless session.

Complexity in Research4Life and Consortial Summon Structure

Low and middle-income countries are not uniform. Amongst the 106 countries, areas, and territories eligible for the Research4Life programmes, there are countries with tiny total economies, but relatively larger average personal annual incomes. There are some with very weak Human Development Index listings but larger total economies, and others with lower development indices but stronger education and information infrastructures. Different countries have different policies encouraging research and differing external funding environments. Each country could be described uniquely, and accordingly, each participating content provider in the programme brings their own perspectives to their engagement with these countries through Research4Life.

The primary motivation for publishers, technology organizations, and others taking part in the Research4Life programmes has always been altruistic. Beyond that, there is also the desire to be recognised for doing the right thing. For some content partners, there is the desire eventually to build markets where appropriate. Some content providers are focused on reaching as broad an audience as possible. Others may be seeking a balance between market growth or sales opportunities and their altruistic goals. Still others may view Research4Life as part of continuum from a tiered pricing structure of “zero” through deeply discounted arrangements offered via other developing country information access initiatives, and finally culminating in the familiar large-consortium arrangements.

All of this diversity in countries and content providers can be accommodated within Research4Life. At present, 38% of the 106 countries, areas, and territories eligible for Research4Life have access to all the content made available by every content provider partner. Many more countries have access to all but one or two of the publishers’ titles. Of the 155 content providers of subscription-based content, 77% provide their information resources in all the Research4Life eligible countries. Beyond these groupings are a large set of free access resources that fit within the subject criteria of the different programmes.

So, for the materials available to 38% of countries and that offered from 77% of publishers of subscription-based literature (and the free access content), it is extremely straightforward to set up the Summon product to indicate what content is available to whom. Where challenges in our work persist are with the 62% of countries and 23% of publishers where access is a matrix as diverse as you may imagine. No two publishers in this subset have made exactly the same choice in terms of which mix of countries they will offer their content in. There are only a handful of countries in this subset where the list of publishers offering to them is the same.

Serials Solutions takes a consortial approach to the task, starting from the top down. To simplify, we will discuss how this works in only one of the Research4Life programmes, HINARI, as the matrix approach is the same for each of the four Research4Life programmes. In this scenario, one main library profile was first established (HINARI). Then, each member country was given their own profile and added as a member underneath the main HINARI profile, in a parent-child type of relationship. This structure allows for general configurations to be made once at the top-level HINARI site, and then automatically be applied to all member libraries. This saves much effort, as each library does not need to be fully configured from scratch. At the same time, it still allows for the flexibility to alter the member profiles as needed.

The parent-child approach is especially useful when it comes to content population. While there are certain resources available to all participating HINARI countries, there are resources where access is country specific. So the first step was to take those resources provided by the 77% of publishers who make their content accessible to all countries and add to it all the open access content to populate the main HINARI profile. This was done using the Serials Solutions knowledgebase (KnowledgeWorks), with resources being added at the provider, database, and title level. Once the applicable resources were added to the HINARI profile they were shared (pushed down) to all the country specific member profiles. Moving forward, those resources will be maintained automatically with regular updates to the knowledgebase.

The next great hurdle was then to populate each country specific member profile with the resources relevant to that country. Considering the number of countries involved (106), this was no small task. Though not as in-depth as a typical Serials Solutions implementation, it is still roughly the equivalent of working on 106 different libraries' profiles. The process involves loading title lists into the member profiles. Depending on the country's holdings, this can be a detail-oriented task requiring a fair amount of time. After the initial implementation though, ongoing updates and maintenance typically require a minimal amount of effort.

The last great challenge is related to access (linking and authentication). Using Summon and 360 Link researchers will be able to discover content and access the full text content on various provider sites. Summon will be freely searchable, but in order to access the non-open access provider content users must be authenticated. Providers authenticate users in different ways, typically by IP address and/or username and password. Since the Research4Life programmes are using a specially developed proxy system to accommodate the complexities of 106 countries, a certain amount of profile specific configuration is needed to ensure everything works smoothly.

Although a fair amount of work is needed to be done initially, the flexibility offered by having independent profiles makes it worth the effort. In the end, each country will have its own Research4Life Summon site. The advantages for the

researcher here are numerous. They will have one interface to search for content, knowing that they are only searching content they will have full text access to online. At the same time, they can still expand their searches to include the entire Summon index, to find even more resources that may only be available as a citation or abstract or borrowed from another library.

Adopting Countries - the Fun of Volunteering

Within Serials Solutions, after the project had been announced and the tasks identified, there was still a challenge to split up the volunteer labor. Since each country would require some level of special set up, it seemed only natural to assign each volunteer a country. The idea of “adopting” a country gave volunteers the option to choose a country they may have already had an interest in and form a more personal connection to the project. For instance, Viet Nam was adopted by a Vietnamese member of the team, and the Plurinational State of Bolivia was chosen by a Spanish speaking volunteer. Further, adopting a country allows for the volunteer contributions to be recognized independently as well as part of the project as a whole.

Though much of the work thus far has been done independently, it has happened in a communal atmosphere. Working in a group setting, the volunteers have had scheduled sessions to all work together on the project. The initial meeting was made into a pizza party where all the volunteers could work together in a more relaxed setting. Though more work is still needed, the “adopt a country” approach has helped the whole group to feel positive about the project at every step.

The Next Horizon

Now we are nearing the end of the initial setup. Soon, the iteration of the country sites will be complete, ensuring that each country sees the right instance when they connect to Summon through the Research4Life programmes. So what are the next logical next steps? A common question during the start-up phase has been “How can librarians in the relevant countries help?” Another has been “My library has access to additional materials beyond Research4Life. How can that additional content be discovered alongside the Research4Life content?”

These are very good questions, and not necessarily easy ones to answer. We will be delighted to have a broader set of volunteers to keep the Research4Life Summon instances updated and expanded to include new materials as they are provided by old and new content partners. However, we are still working through the challenges of how to coordinate a network of engaged volunteers. These challenges include issues of communication as well as information extraction, training on back-end interfaces, and much more.

The good news is that once implemented a Summon site can remain live and stable with minimal effort. The Software as a Service (SaaS) model is used, so no hardware will need to be maintained locally. The Summon interface is constantly being improved and updated, and new content is continuously being added to the index. However, a country's holdings are apt to change over time, with new publisher partners joining the Research4Life programmes, new titles being launched, old titles being removed, coverage dates changing, etc. So this is where the most basic need for help will lie.

To put it into perspective, Serials Solutions products are generally maintained locally by librarians. On a basic level, staff time is often spent managing library collections—working with the knowledgebase to keep a holdings profile accurate and current. Depending on the size and intent of the library, maintenance may require only a few hours a month or several librarians managing separate modules on a daily or weekly basis.

The HINARI sites will likely fall on the low side in regard to ongoing maintenance. Of course, if needs and resources were to permit, having someone locally who can consistently maintain and optimize the profile is certainly to the benefit of the users. As librarians are familiar with their unique user groups, they can easily customize their sites to suit their needs. For instance, working together, a country's librarians may choose to add public notes to inform their users of the log in requirement prior to accessing an article link. Many elements of the Summon site can be customized as well, such as changing the available search facets and, while not as relevant to Research4Life, the site logo. Some training is certainly useful, but most customizations can be made in a user-friendly administration console and do not require specialized computer skills.

The question of whether it will be possible to define additional sets of content beyond Research4Life resources for discovery in individual libraries is a further challenge. As the reader has seen, Research4Life by itself is a complex set of resource allocations and rights. It is a little difficult to imagine at this stage adding yet more complexity to the picture. However, the situation is not a unique one. Imagine a library in Finland which belongs to two different consortia. Such a library would want to inherit Summon instance rights from two different parents. This would require further work and engagement from volunteers to help ensure the proper sets of content are accessible.

Another way local libraries can help is to encourage more providers to agree to have their content indexed for discovery. This is particularly important with local publishers in low and middle-income countries who may not yet be aware of the possibilities of having their publications more widely searchable.

For now, we are delighted by how far we have come along the path of making information retrieval so much easier for the novice and casual information users in low- and middle-income countries. And it is not just the casual user who will

benefit. We are delighted that we can also provide a boost to the explorations of more experienced and in-depth searchers as well. We can look forward together to building on this work for wherever our next steps may take us. While the scope of this project is already large, it can certainly grow and evolve in any number of ways. Besides HINARI, Research4Life includes three other programmes: AGORA, OARE, and ARDI. Each will require a similar amount of effort to implement as HINARI.

The project is only just beginning but ongoing research can be done to analyze the usage of the sites and measure the impact of the project. Perhaps new insight can be gained about usability in the developing world. We are excited by the potential opportunities and we look forward to the many new developments and connections that we can only see the outlines of for the moment.