April 2015 | Whitepaper

Graham Beastall
Managing Director, Soutron Limited
graham.beastall@soutron.com

Search Portal Design

CHALLENGES & OPTIMISING IMPLEMENTATION

Soutron

Digital Knowledge Services

SOUTRON'S BACKGROUND

Soutron was founded in 1989 by Graham Beastall and specialises in digital library technologies.

Between 1991 and 2012, Soutron sold systems on behalf of third party software suppliers, to special and corporate libraries. In 2007 it was clear that suppliers were content to rest on legacy code and avoid new investment decisions. Soutron saw this to be an opportunity to create new systems and provide a complete service to clients.

Soutron is committed to develop a new form of library management system with the flexibility to address different needs and concerns of all libraries and knowledge centres. Our goal is to make processes simple to follow and software easy to use for multiple libraries within a single global system.

The Soutron system is robust, proven and in use by over 200 organisations in the UK, Ireland, Canada, USA, Italy, Switzerland, Sweden and Belgium across a variety of sectors.

Clients include large and small organisations that seek to manage their data in a more controlled way with the benefits set out above.

Graham Beastall has worked with libraries and global organisations for more than 30 years to devise systems and solutions to satisfy information management ambitions.

The information provided in this document is confidential and provided in confidence for the sole use and benefit of the intended recipient and is not to be physically or electronically copied, shared, transmitted or the content communicated to any person or organisation without the advance written permission of Soutron Limited.

CONTENTS

SEARCH PORTAL | AN INTRODUCTION

‘Search’ is an important part of any database solution. In a Library or Information Management System, the traditional ‘online public access catalogue’ (OPAC) was designed for bibliographic data - data that is generally well understood, without the need for security or controls determining who can access it. In recent years, more diverse content from internal and external sources has led to the introduction of Discovery Services - a single index originally targeted to University establishments. The Search Portal has broader ambitions than both OPAC and Discovery Services, being a one-stop service for end users in an organisation to locate data to support their work.

Planning and implementing a Search Portal is a challenging process, with many aspects to consider. A Search Portal has to fulfil a number of objectives and is the ‘shopfront’ for information services. This paper presents practical tips from our experiences across several industry sectors and might stimulate ideas. It is not prescriptive or formulaic.

A Search Portal has to fulfil a number of objectives and is the ‘shopfront’ for information services.

A Search Portal requires considerable planning and analysis of the many users’ needs that it will have to satisfy. Before any attempt is made to configure a solution, a survey of user needs and data sources is recommended. A library catalogue with its bibliographic data is relatively simple to set up compared to a Search Portal. The extended reach that a Search Portal offers plus the wide range of content made accessible from disparate sources places a much greater challenge on those responsible for design and implementation. It requires knowledge of library science, knowledge of the way content is structured and involvement with users.

The scope of the Search Portal is much greater than a catalogue and must bring together the library catalogue and different databases, so the end user has one place to navigate through the variety of data. Presented as a gateway, the Search Portal removes the need for users to login to different systems, improves the user experience and the quality of the results returned.

THE NEED OF THE USER

There are many reasons for an end user to reach for a Search function. Users may want to find: a specific reference or document, a textbook or a table within a particular book, an argument based on law, a digital asset collected by the organisation or information in the form of 'know how' to answer a question or to help progress a decision. Instead of reaching for Google, the organisation invests in creating a service to respond to such requests.

If we ask end users what they want from a Search Portal, they will often default to what they already know. The Google-like option is cited as the preferred type or style of search. Those that are time poor want a quick and easy to use solution. The Search Portal project team may wish to form a focus group of end users and introduce them to the many additional search options available. These search options may appear more involved than Google, but will in fact save time in the long run by returning more accurate and relevant results. Users may also find that what they initially thought to be a simple search might, in fact, turn out to be very complex. Assistance from a Librarian may not be readily available, so thought needs to be put into how the Search Portal will work for the end user without a librarian being on hand. Running training sessions and creating user handbooks will assist in the psychological and operational shift towards more in depth search criteria.

The way that results are presented can be customised through sorting and search algorithms to rank results. Relevance ranking might be helpful if the ranking can be customised to suit particular content. Results are delivered based on quality over quantity, with fewer results of a close match being favoured over numerous results that are tenuously linked. The objective is to find the content that is most relevant rather than finding thousands of loosely matched results. When a user retrieves a set of results, if the number of hits extends beyond two pages, they are unlikely to navigate beyond the second page. A Search Portal therefore needs careful planning, analysis and development. It should bring about the following benefits:

- **Bring together information – a one stop service**
  Multiple types of content should be included in a unified search result.

- **Provide a 'no fuss' user experience**
  The results should be presented in a clear format, within a familiar interface that allows the most relevant result to be quickly identified.

- **Enable targeted results**
  A Search Portal should allow the user to focus in on just those results that are most relevant, perhaps appropriate to: a subject, location, type of material or the area of practice or business that the user is engaged in.

BUILDING THE PORTAL

A number of constructs are key to creating a Search Portal. These are common across all Search Portals but the emphasis may vary depending on the complexity of the data and the needs of the organisation. If the industry being served only has electronic resources stored outside of the organisation, it may guide the process in a different way than if the data is held internally.

Experience shows that when setting up a Search Portal, careful attention to the following four areas is recommended.

ANALYSIS & PLANING	**DATA**
CONNECTORS	**SIMPLICITY**

The task of building a Search Portal starts with analysis.

ANALYSIS & PLANNING

'Search' is not like any other workflow process. It seems a deceptively simple task, yet as Google has proved it takes millions of dollars and considerable computing power to put together what is a simple search. We strongly recommend devoting time to the careful analysis of your business needs. Some commentators have described this phase as being like gardening - a plan is needed. It should incorporate the direction of the sun, the type of soil, the irrigation system, wind conditions, height above sea level, particular features that are wanted in the garden etc. It needs an interview with whoever the garden is going to be created for to understand their motivation and how the individual might wish to maintain the garden. Sowing plants is not enough. It also involves cutting, feeding, weeding and watering continuously. Search functions are similar. There are many elements that need to be taken into consideration in order to have an easy to manage and popular search facility that meets the needs of the end

user. Once built, if it is left without being tended or cared for, like a garden, it will soon fall into disrepair.

A project plan, setting out the various tasks and indicating resources needed to implement the Search Portal, is a useful document to assemble. It is a focus point to think through what is needed, how it will be assembled, tested and loaded with data and made ready. Placing realistic timescales on the tasks will help put a perspective on when it will be ready to test and release. This informs not only users but also more importantly managers, IT and other personnel who will have an input to the system delivery.

When creating a project plan, allow sufficient time for identifying all steps of the process. If this phase is not undertaken then the Search Portal will most likely fail to achieve its purpose. There may be several people to consult and verify requirements, suppliers to speak to, data sources to examine and measure, and test scripts to create and validate. Many different aspects affect search: content, metadata, user experience, ranking, sorting and performance. All have to be maintained over time and continuously updated to ensure search results remain relevant, accurate and worthwhile to the user.

User observation, usage mapping and identifying frustrations will guide the way your solution is constructed.

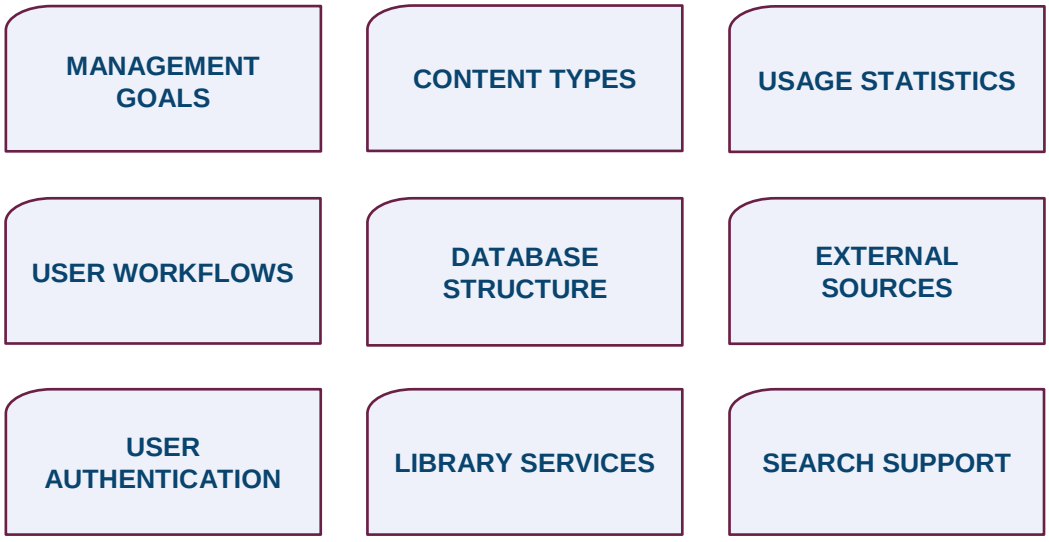
There needs to be a clear understanding of the context in which the user is operating. Understanding their motivation is critical. In a legal firm there is likely to be a different motivation to that in a government department or in an archival service. Observation may be needed of the search tasks that are undertaken by a user. This should extend to how information found is used and within what timeline. Online surveys may provide one form of data collection, but may not tell the whole story. Focus groups and one-to-one sessions with users are great ways to observe and learn behaviours. Build up a matrix of users and user types or groups. This will help classify their data usage requirements and the key elements needed by each. This helps determine if one homepage is sufficient to service all users, or if more than one interface design is needed to cater for different user groups. Find out what data sources are used today and try to find out why and what level of satisfaction exists.

The type of searches that the user makes and their frequency may be related to the specific job being performed. This may mean that the metadata and database structure has to include certain levels of detail to help the user get the right results. A change to the way in which data is collected and indexed may be needed. User observation, usage mapping and identifying frustrations will help to guide the way the solution is constructed.

If one of the goals of the Search Portal is to reduce the pressures on library staff by enabling users to serve themselves, this insight could be invaluable in shaping the Search Portal. The portal could be configured with pre-defined searches to save the users having to construct their own search strategies. Some complex searches may need library expertise to construct or may need significant processing time to deliver the results because of complex search syntax. In such situations it may be necessary to call upon the supplier of the software to help create advanced searches that can be pre-processed or crafted by a programmer with skills that are either not available to the information staff or require direct access to the database tables.

It is recommended to have several examples of the types of searches that are being conducted, as these can then be used to create scripts for testing later on in the process in order to measure the results against what the user might expect now. Classifying these according to the type of user, the type of search, the degree of complexity and the likely quality of results will help when introducing users to the new Search Portal.

When analysing requirements, consider these nine aspects:



Setting down in writing the goals of the Search Portal is worthwhile so that all associated with the development of this new tool can be reminded what the end result should be. It provides a reference point throughout the project. A goal might also encompass a resource statement or a target that describes a measure of success.

Metadata analysis requires the ability to match data structures and search needs.

DATA

There are a lot of questions to ask about data:

- Make a list of all databases, systems or repositories that are likely to hold the necessary data for users to do their job.
- Document the types of users that the Search Portal will have to service.

Analyse the two and set up a matrix to see where commonality occurs.

- Does the library manage the data sources in-house or are they located elsewhere, external to the organisation?
- If multiple sources of data are used, is there any level of duplication?
- Is there potential for de-duplication to be built into the Search Portal?
- Is it possible to filter out duplicate data in advance by examining the sources and determining one source over the other based on a quality assessment?
- Is it essential that data sources be accessed in real-time or is data less time sensitive?

These are some of the questions that might be posed when looking at data for any information service. Set up your own list in a spreadsheet, maintain it throughout the project and update it post project with results from periodic testing of the system. Include descriptive elements to make information easy to find. These descriptive elements are called metadata. For example, a tin of baked beans is described with the name of the brand, the manufacturer, the weight, the ingredients etc., these are all metadata. Whether the information being searched for is about an article (printed or electronic), a document (PDF) or a book (eBook) it is the metadata and the quality of that data that determines the success of the search results. This makes finding the data simple and more accurate for the end user. The use of controlled vocabularies, validation lists and field validation is not only helpful in making sure that the metadata is consistent, but also allows the user to enter variations into search statements and still find relevant content. In order to determine the quality of metadata, it is recommended that:

- Staff examine samples of metadata records, comparing them to typical searches for the type of content that it aims to describe, to reveal potential gaps or inaccuracies.
- A small sample of end users are selected to complete pre-defined tests using the metadata to ensure the search results are accurate.
- Taxonomies are verified to validate vocabularies and descriptions. These may require a significant amount of work before being ready to release to users.

Where a third party system holds the metadata, it is suggested that searches are made directly on the source data to examine the results. Measure the consistency of results over a period of time. In the Search Portal a user may wish to search a single source of data or several all at once. Where the internal and external data is searched simultaneously, the results might be displayed in the Search Portal in separate tabs or windows. This could allow the results to display in the native application.

A Federated Search engine will combine results, from a single search across several repositories of data, into a single relevance ranked display. This is a challenging task to take on as a piece of development. Without use of a commercial federated search tool specialist, programming skills will be needed together with continual maintenance of the tool. For this reason alone if you need real-time searching across databases, analyse the data results that come from such a tool.

A set of typical searches that a user might wish to perform, will inform and help determine the best way that the metadata is structured. If further classification or indexing of the data is required, in order to deliver the appropriate results, this should be resourced, as it will have a positive impact on search results and user experience. Analysis of the metadata is incredibly important. Unless the structure and the content of the metadata is examined, in the context of search strategies and the way in which users expect results to be presented, the result will always be a poor experience for the end user and will lead to ultimate dissatisfaction with the system. The search function may perform well but if the metadata itself is incomplete or is poorly constructed, no search technology will help. Metadata analysis requires conceptualising a search requirement against a data structure. If the skill does not exist within your team it's worthwhile getting someone with expertise that can perform this type of analysis. A web master who is great at putting a system in place may need specialist help to assist with determining the metadata suitability. Often more effort needs to be put into building taxonomies or facets, to allow a user to easily explore the content, than is first expected.

The assessment of metadata calls for honesty. Search success depends on it. If the metadata is not consistent or it is poorly structured, so that it cannot be easily interpreted and displayed in search results, then there may need to be work carried out to improve or convert the metadata into a more appropriate structure to achieve the search results or the presentation style that is required. Metadata in some cataloguing systems has poorly defined structures but is not immediately noticeable because programming in the native system shields any inadequacies. The effort to address structural issues with metadata is usually worthwhile, simply because the user can quickly assimilate the results and say, "great, that's what I need".

CONNECTORS

The selection of data sources is paramount in the success of the Search Portal, but how the data is accessed and presented can differ. Data may be available in a pre-prepared index that is maintained and supplied as a Discovery Service, a very large index may be pre-built and configured. This type of data source is very useful in that it is fast to search and has a large set of data, from a variety of sources that can be filtered and offered to users quickly and easily. The alternative is a Federated Search where database sources are queried directly in real-time and results returned. The ideal might be a solution that combines both, since some data sources will not allow their data to be loaded into a Discovery Service Index. Discovery Services, by their very nature, are continually changing and care should be taken to select the Discovery Service that matches your data needs.

The ‘connector’ is a program that sits between the Search Portal and an external database. It is unseen by the user and works in the background to translate the search statement into a syntax understood by the external database. When results are returned by the external database, the connector transforms the results into a format that can be displayed to the end user in the Search Portal. Each external database will need its own connector. A new connector can take between 2 to 3 days to create, depending on the complexity. External sources might change their syntax without alerting anyone. A subtle change can affect the results served, or may break the connection altogether, resulting in the need for maintenance of the connector. So any solution dependent on connectors needs to be monitored and maintained.

Connectors are required to manage:

- **Multiple content types**
  Different content sources might contain different types of content. It should be possible to aggregate these. The connectors need to be able to reach external systems. Filters and sorting tools are required to manage returned results.

- **Security models and permissions**
  A security model for different types of content may need to be implemented. Security is one of the biggest challenges facing Search because it demands extra time to serve the results.

> **Achieving simplicity is both complex and difficult.**

- **Metadata**
  Metadata is the finding agent. Without good metadata the search will be less effective in producing relevant results and each type of content may have different metadata structures. This is used to display, filter, sort or refine the results.

- **Advanced features**
  Each business area is different. For example, managing skills and expertise in a database would require a very different set of search and filter criteria to that needed for articles or books. Identify these differences in the analysis phase.

The use of connectors is one approach to access multiple data sources from a single search. A Discovery Search Index may be preferred and is one that can be licenced and embedded seamlessly into the Search Portal. Speed and performance may dictate the path to take, as the perception of the speed of results may be a critical factor in end user acceptance.

SIMPLICITY

Achieving simplicity is complex and difficult, requiring more than attention to design and layout although this is very significant. It relies on analysing the needs of the user community and having done all the other suggested tasks before getting to system interface design. There may exist more than one user community and it may be helpful to create more than one Search Portal so that each user community is able to have their needs fully addressed without compromise. This type of requirement will surface when a detailed analysis is conducted. By approaching the project goals in this way it may ultimately simplify the Search Portal design.

The cleaner and simpler the Search Portal, the better the user perception. Terminology needs careful assessment. A term may be familiar to information staff but not to other users of the system. Words and text used in the Search Portal, therefore, need testing to ensure that they convey accurate meaning and to ensure good practice. Choosing how to describe actions and tasks can expand the use of the Search Portal. Rather than offer lots of search boxes to complete, for example, it may be better to offer a link to a predefined search that lists results that the user is then able to filter.

SEARCH PORTAL DESIGN

There is a natural urge to get the appearance right. It is not unusual for many people to be involved in this part of the project and decisions stall as a result. Prepare for the fact that someone will not like what is being proposed. Change itself will impact judgement. Where to start the design process? There are different approaches to user interface design. Some that we have seen but would suggest avoiding are:

- Cram as many links and options on to the first page as possible
- Limit the page to offer Search and no more (similar to Google)
- Emphasise the brand at the expense of services
- Use multiple fonts and styles and make the user look hard for what they need
- Extensive help text that creates confusion more than clarity

> **The cleaner and simpler the Search Portal, the better the user perception.**

We would recommend thinking about what each type of user wants from the Portal and concentrate on their priorities. The look and feel of the user interface is likely to be commented on by everyone that is exposed to it and it becomes very subjective. So base designs on research and focus groups where a few representative people of the user community are involved. Test assumptions and prepare mock-ups of the layout to check you have all the essential elements.

Commercial online databases may have interface designs that have become de facto standards within their particular market segments. The end user, who is exposed to these, may establish familiarity with a type of searching. Replicating and adapting the existing interface design of that database provider is one path to pursue should it deliver a good user experience. It may, however, be very specific to a particular type of content and may not be applicable to other types of content. The benefit of maintaining a certain look and feel is that it will minimise the requirement for training sessions due to the similarities between the systems. Commercial databases, however, often change search syntax in subtle ways and it may not be possible to emulate such systems without extensive programming.

Pre-prepared templates (if available with the Search Portal software) can help the designer to formulate ideas about what to include, layout, fonts etc. This eases the way into crafting a superior interface. If you don’t have a company brand to apply to your Portal, using a template also means you can avoid investigating the possible fonts, images and layouts that are available, including such things as which fonts are web safe and easy to read. Performance can be affected significantly by the choice of fonts, image sizes, the number of controls on the page and how the page is assembled.

DESIGNING THE INTERFACE

A pleasing user experience is as important as having good metadata. Good layout and positioning is helpful to the process of finding information fast. Too many distractions on a web page are often a major barrier to acceptance. Six aspects to explore are:

- **Search**

  A search box is the normal way to initiate search. This is usually a facility that performs a search across a number of fields (data elements) so that the user does not have to think about structuring a search. The search is usually configurable which means fields can be included or excluded so the search only looks for information in certain places and not others.

  If the search is made using a phrase such as ‘decision making for financiers’ rather than individual words, this may result in users getting sub-standard search results, unless the Search function can distinguish between a phrase based search and a keyword based search. Check the list of words in the database that are NOT normally indexed as it may make searching easier by incorporating these into the index that is used. Alternatively, some very common words might be best not indexed and these should be added to the list of ‘stop’ words.

  A quick link or a pre-defined search is an alternative option to include. It is executed in real-time to show all results that meet the pre-determined search criteria. This is distinct from presenting a link to a set of pre-defined results. This is called a collection or a ‘saved set’ and is another way of presenting data without the user having to enter a search strategy.

  There are generally two forms of search: simple and advanced. By having separate pages for each type of search, the user is helped either by having distractions removed, or by being presenting with extra search options to enable a more accurate search. The advanced search options should be tailored to the type of content. It should be obvious and easy to move between simple and advanced.

- **Filters**

  By knowing your audience and what form the data content will take, may mean that ‘filters’ can be presented to the user, allowing quick selection of terms to increase or reduce the number of results. The added option to remove words in

> **A pleasing user experience is as important as having good metadata.**

the search means a user can control the results, without having to know that much about search syntax.

The option to select filters could be displayed as part of the initial search criteria to limit the content retrieved, or could be shown with the results, post search execution, to focus in on the most relevant content. For example, a user could refine a search by selecting or de-selecting a term or description to expand or restrict the search results. This is particularly important for specific practice areas or business functions where the analysis of needs has revealed particular ways in which information should be sorted and presented. Some data structures may be too large to allow simple display of a set of filters. If there is a large taxonomy or thesaurus, this may be better presented in a different way by limiting the number of levels available or by pre-selecting particular terms to build a set of selectors.

- **Result views**

  The control over the way results appear is such a significant aspect that we could spend a whole paper examining how best to do so. Should they be presented with a full record display i.e. all fields? Or with a brief display of the most important data that helps distinguish one record from another? If multiple types of content are present, a particular type of display format might be needed and how that can be set up may require extra capabilities within the application. Or, if the application does not have these, it may require the use of an API or web page for customisation. This all needs thought and careful usability consideration. Reducing the number of clicks to get to the important content and serving the best results on page one or two will optimise the user experience.

  The software that has been selected may have the capability to group content or to present it in separate formats. Prepare a set of sample records for each type of content and hold a focus group asking people to comment on how they think the metadata should be presented. A discussion of the presentation of results may also lead into how the user can easily transfer them into another application or paste them into a report or share them by email.

  The sort sequence and the ability to set a rank in terms of relevance are also important aspects of the results design. It is useful to refer back to the observations that were made of what users were looking for and how they wanted to see data. Anything that brings the most relevant content to the fore,

reduces the number of pages of results or decreases the number of mouse clicks required to lead a user through their search journey will get positive feedback.

- **Layout**

  Think of the page that first greets the user in the same way someone going into a house for the first time might feel. If there is a huge amount of clutter and irrelevant options the user will be districted. The better the user experience, the more it can help users make decisions. Sometimes, displaying information in a traditional way is not enough. The faster the user can digest what is on screen, the faster their work is completed.

  White space is a valuable tool to help the user and is often overlooked. It is important to understand how best to use it as it supports visually finding the content on the screen.

- **Mobile**

  Where is the user most likely to use Search? Is there a need to work with the data outside of the office or away from the desk? This may also be a consideration – do mobile devices now come into play? Is the screen layout designed to allow the user to get an answer when on the move? By having a HTML5 compliant Search Portal, mobile optimisation can be built into the interface, enabling the page content to be displayed according to the screen size. This is becoming an expectation as more and more users turn to mobile devices and are accessing content on the move, out of the office.

- **Related Records**

  When users are served search results, it may be useful to display other related records, together with any other connected content. This way, users can quickly navigate to other potentially useful information. The Search Portal and database software may not have such relational capabilities but it is worth asking how you might achieve this, as it provides extra dimensions to presenting results. It is especially useful in archival circumstances where there may be library, images and archives all stored in a single database.

> **Search is incredibly important.**
>
> **A Search Portal is a strategic tool that helps reach out to users.**

MEASUREMENT

A lot of the work in setting up the Search Portal focuses on the content and design. But the metrics arising from use of the Search Portal is not to be overlooked or excluded from consideration. Think early on about what information is needed to measure the impact or success of the Search Portal. Assessment may need to go as deep as knowing who is adding metadata and uploading content. The metrics that can be collected from the Search Portal informs how it can expand and grow, how it should be maintained and nurtured. Define the level of reporting that is needed and ensure that the database and Search Portal has these usage statistics built in and can be reported against. If the content incorporates published articles, usage statistics may be built into the product and these can be obtained in the form of Counter or SUSHI statistics.

FINAL THOUGHTS

Search is incredibly important. A Search Portal is a strategic tool that helps reach out to users and can be applied in a variety of ways. It requires a commitment to make it work and so once established, it is worth devoting time to analyse this aspect of the information service every week. This includes looking at and monitoring usage statistics as well as monitoring content and what is being consumed.

Consider how your organisation can maximise the use of content and encourage users to access qualified sources of content by presenting the Search function in the most appealing way. It may be that you can encourage other information providers in the organisation to present their data via the Search Portal or offer clients access to materials as part of your service.

Complementary training or open day education services to support your users' Search activity can be very helpful. Where we have held open days for clients, to support the introduction of a new Search Portal, we have often been told that the very act of notifying users that they are welcomed and supported in their Search has had beneficial effects. Incorporating Live Chat to assist users via a message service can also be a good way to provide added service support providing your other duties are not adversely impacted.

Advertising library services on the Search Portal web pages could be a valuable way of expressing the added extras available to a user. The services may be as important as the Search Portal itself in the long term. Consider what, where and how the Library and Knowledge services might be promoted within the Search Portal.

HERE TO HELP

Soutron are here to help, advise and provide assistance. Every organisation is different and unique. Let us know your experiences and needs so that we can continually improve systems for you and your users.

If you have specific needs, or have created Search pages that you would like to share with us, we'd be very happy to hear from you and share the good ideas around our user community.

Points of contact

UK and Europe

Soutron Limited
Highgate House
Burley Hill
Derby DE22 2ET

T: + 44 (0)1332 844 030
info@soutron.com
www.soutron.com

USA

SoutronGlobal
1042 N. El Camino Real
Suite B-215 Encinitas
CA 92024

T: + 1 (0)760 870 4243
info@soutronglobal.com
www.soutronglobal.com