Search: Considerations for Running an Effective Pilot

Proof of Concept > Pilot > Deployment for Enterprise Search

Deployment and adoption of Enterprise Search technology does not necessarily follow a standard Proof of Concept > Pilot > Broader Deployment approach as with other legal technologies. Deciding which enterprise search system to implement and partner with at your firm is almost always taken before the system is installed and linked up to any of your data. It is a huge job to clear Security, get the environment installed, and your content indexed, so a proof of concept or 'mini-pilot' can be difficult for either party to agree to.

A more standard approach is to treat that initial proof of concept as the initial requirements and evaluation discussion instead. See the Evaluation Framework and the Requirements Document for the types of discussions you'll be having during this initial phase – and possibly with more than one vendor. This is a useful step to also see how your future relationship with the vendor is likely to go. Looking at your requirements and working out which are your 'Must-Have' and 'Deal-Breaker' requirements (and why) are fabulous initial discussions to have whilst looking at their demo versions and screens.

Reach out to your vendor to understand the successful pilots they have experienced in the past. And reach out to the industry and other clients to understand what a successful pilot looked like for them.

As you work through which source systems you are looking to index and display as 'search tabs', you can identify one of the smaller collections to index first as your actual 'Pilot'. Assuming all goes as smoothly as can be expected with a system this large, then this pilot roll-out becomes your first phase on an ongoing deployment plan. Choose a smaller database with a specific set of users that is a mix of search specialists, content owners and lawyers to give you the best range of feedback in this initial phase.

Another deciding factor in how this first phase will go, is whether you already have an enterprise search engine or another way of searching multiple collections. If so, you may be doing a side-by-side comparison as part of this pilot. This can initially be useful to set minimum requirement levels to ensure there is no drop in search experience for your users. However, there are advantages and disadvantages that are not easily comparable on a like-for-like basis so be careful it doesn't lead you astray.

Search Prep

If you do have a search engine – whether that's just within your DM or whether it's another Enterprise Search vendor – one of the most important steps to take in preparation for a pilot is to baseline and explore your search data. Pulling your search logs plus usage metrics (from the prior 12 months at least) will give you the most definitive insights for understanding your users and their needs. This data will give you the specific words that people use to search. It will also show you your specific 'Search Demand Curve' (see: Search Analytics for Your Site – Louis Rosenfeld (Rosenfeld Media, 2011)).

You will see the "Fat Head" of small number of popular search queries at your firm, as well as the insane number of unique queries that sit in the "Long Tail" of the data. This will help with the development of your search testing plan so you can repeat some of these more frequent searches to review result sets. It will show you how advanced your users are in their searching methods and behaviours. It can show you how many searches are abandoned, and how many 'zero results' journeys happen at your firm. Plus it can show you content that is missing, as well as show you who might be good potential pilot users. And by baselining your numbers you have a way of measuring success of your deployment in the long-term.

1. Define Pilot Objective

To begin, you need to identify what you want to specifically test for. You will already have an idea of what search will look like "out of the box" from your initial discussions with the vendors. The decisions you now need to make are what will this look like with *your* content indexed. You will need to agree an evaluation feedback loop with your users, your IT team and the vendor to make tweaks to the data and the UI during this phase.

In this initial phase you will be diving deep into the data, metadata and taxonomies in your databases and working out how to align them to the user interface so that it is useful, usable and engaging.

2. Evaluation

Having identified what you are specifically testing for in this pilot phase with your Must-Haves and your Deal-Breakers, create an evaluation questionnaire for your pilot group of users to use and track their scores against. You may have one for end users and one for your project team to capture different types of feedback.

For example, end users can test for things like success of search, 'ease-of-use' and speed of results with a scoring system for overall experience. Whereas your expert users and project team may be testing for specific functionality, such as: application of multiple facets to a single search, printing of results screens, review of data quality in the filters and snippets (and not just on the first page of results) etc.

A testing plan with specific searches can help ensure more consistent and comprehensive testing and the results found for different users. You should have a list of search terms that can test ethical walls, as well as those that can reveal content that may have incorrect permissions applied (eg, 'salary' and associate reviews).

Remember to capture quotes and testimonials from your pilot users from demos and feedback sessions that can be re-used in future firm communications and roll-out plans.

3. Data + Content

It will be useful to create a spreadsheet to capture your different metadata choices for the deployment. For example, your columns might be: Field Name, Field Label (for UI) and any formatting requirements for that field (eg FirstName LastName), a description of the field (along with whether any fields should be concatenated in the UI – such as client ID and client name), whether it should be indexed for search, whether the field should be used as a facet/filter, whether it should be included in the search snippet, whether it should appear on a Preview Screen (eg. in a metadata panel) and whether it should appear on the Advanced Search page etc. You will need a sheet for each search tab you are planning for.

Looking at a single collection or search tab, you will need to plan which filters you will have down the left-hand side to help lawyers narrow their search. This will largely be dependent on the mandatory fields or auto-system fields that are completed when saving to the database. The quality of metadata applied to a document when being saved to your DM (or other content management system) will dictate the quality of the filter used when searching for that document again. All of the data hygiene and quality will be exposed really quickly during this first phase of your project. And, to warn you, some of it will Not Be Nice.

Here are three different metadata fields to use as examples of the kinds of questions to consider during your pilot.

Author

- This should be largely system-automated.
- How is the author name captured in the database or source system?
- How do you want that formatted on the UI? (ie. will you be pulling the system name from the database directly (eg. simpsonk) or will you want to re-format it as Simpson, Kate or Kate Simpson?)
- How do people refer to each other at your firm today? By first name or last name?
- If documents and emails are being returned in the same tab then you will have to deal with a similar list of 'Senders' – ensuring you have a consistent and standard format will help with retrievability on a Search UI
- Also consider future tabs for how your internal lawyers names might be formatted and how your databases will have different structures (such as Billing Lawyer in your Billing/Practice Management system)
- Plus: are there multiple versions of a single author's name in the database (K. Simpson and Kate Simpson) that will need normalizing and mapping to a standard?

Document Type

- Is there a document type field that can be derived from folder metadata in your DM (eg. if one of your DM folders is called Agreements, can all the documents in that folder auto-apply that metadata at the document level?)
- How many of your documents and emails have "Admin" as the document type because it is the first category in the drop-down?
- Does the prevalence of this potentially poor data quality issue mean that this filter should not be used or is it "good enough" for initial purposes (and possibly illustrative purposes to improve future tagging)?
- Do you have any auto-classification tools that might help fix this issue or any batch changes that can be made to the system prior to indexing to resolve some of this data quality?

- Consider having an 'Application Type' or 'File Format' as a useful alternative category to have in addition to document type on the search UI
- Be aware of the labels used in a database and the labels that you will want to display to your end users – should it be .doc, .docx or Word or MS Word?
- Will you want to display the MS Word icon to help with easy identification in the search results somewhere? Do you have recognisable icons for all your different document types and file types etc?

Practice Group

- How is the lawyer's practice group derived in your DM?
- Is it from the author and mapped from AD (or people database), or downstream from the matter information?
- Do you have multiple practice groups for each lawyer and allow for multiple tagging in your Client/Matter system?
- How will this affect the quality of this field to be used as a filter?
- Do you have any auto-classification tools that might help fix this issue or any batch changes that can be made to the system prior to indexing to resolve some of this data quality?

4. Communication

Consider your preferred method for communicating changes you want made to the UI with your IT team, any third party professional services, and the vendor. Using a screenshot tool like Snagit can save so much time in visually describing the changes that you'd like to see during your testing phase. Schedule regular testing feedback loops with your users and summarise that feedback and those requests into your visuals.

5. Training and Support

Gain access to your vendor's user or knowledge base so you can start connecting with their ongoing support protocols, as well as identify and access documentation you can reuse and repurpose internally. For example, understand the search operators and wildcards used in the engine so that it can be turned into a quick 'Searching Tips' document for your pilot users to refer to.

5. Time Period

Your pilot should be a set and focused period of time. We recommend this is no shorter than 3 months. This seems like a long timeline but if you want to see updates made in response to user feedback you will need to build in time for those changes to be gathered and made (if possible). This process will give you a feel for how scalable and flexible the system is.