Searching for answers

Written by Stephen Arnold → sa@arnoldit.com

Implementing enterprise search software can be challenging, but following a few basic rules can ease the process, says consultant Stephen Arnold.

E nterprise search and retrieval is a tricky field. Due to the wide variation in the type of information sets and business models within and among organisations, there is no single recipe that works every time.

However, there are some basic rules to follow. Indeed, the larger the organisation and the greater the number of variables, the more important it is for the search implementation team to work in a methodical, systematic manner.

The following checklist should cover each of the main bases:

1. Storage

This refers to the amount of storage required to house an index, the document collection if a repository is implemented as part of the search system, as well as additional capacity required to hold temporary index files and any other "overhead" needed by the system. This is required to perform such tasks as classification, to generate reports, and to process documents in multiple languages. The importance of storage cannot be underestimated. It can be a critical "gotcha" in many search projects.

2. Servers and workstations

Separate servers may be required to handle such functions as administration, maintenance and performance tuning. Large-scale systems for indexing millions of documents require dozens of multiprocessor servers, workstations, storage devices and other machines.

For 24x7 availability, systems may need a mirrored architecture with redundancy built-in to each component. If it is being implemented in an organisation with stringent engineering practices, it will require its own development, staging and production environment.

3. Staff

Search systems, even the appliances offered by Google and Thunderstone, require staff to install, configure and monitor them. The more elaborate or sophisticated the search system, the greater the number, type and roles of the individuals required to perform necessary services.

It is not easy to generalise about how many specialists is the "right number" to employ on an enterprise search project, particularly as some will require the user to build the search system from building blocks provided by the vendor, while others are designed to operate with little or no support.

The best way to estimate the number of engineers or support staff required is to collect data and analyse that data before signing the contract with the search system vendor.

4. Management support

Early in the process, the project team should ensure that management supports the project wholeheartedly as an ongoing concern. Enterprise search systems always require some tuning after implementation, but many end-users will have high expectations.

When the new system is finally introduced, some end-users may express their disappointment loudly and aggressively. Without management support, the enterprise search deployment can stall at that stage.

Part of the management support effort will include a plan for the project communications about it. That will require the identification of a project owner as well as owners of key sub-tasks in the project. Without management support and its associated "search plan," the risks associated with the search system will rise substantially.

5. Budget and cost tracking

Regardless of the size of the organisation, an enterprise search project – like any other – will benefit from a budget and a formal cost tracking mechanism. But cost overruns go hand-in-hand with enterprise search for a multitude of reasons, not least the "learning curve" associated with any software installation from the moment the first code is loaded onto a server.

6. Timetable

A timeline for a search system deployment is a necessity. But the schedule should not begin with the installation of search software, it should:
• commence with the pre-selection specification;
• include a head-to-head "bake off" among likely vendors' products;
• reflect the construction or upgrade of the infrastructure;
• account for critical up front document analysis work.

The timetable must include firm deadlines, as well as contingencies for debugging, the writing scripts to work around unexpected issues and staff training staff.

Some vendors talk about a one-day deployment. This may be true for outsourced and hosted solutions in some circumstances, but the planning and preparation alone typically requires weeks or months of work. **M**