

Searching for Search:

A High Level Guide to
Finding the Right
Information Access
Solution

Interview with US CEO at
Exalead. Paul Doscher

Paul Doscher joined Exalead as the CEO for North America in 2008. A software veteran in the IT industry, Mr. Doscher was most recently CEO of Jaspersoft, where he led the company to become the worldwide leader in the open source business intelligence market. Mr. Doscher has also held executive management positions with leading high-tech companies including VMware and Oracle.



According to surveys, most companies are dissatisfied with their ability to easily find information inside their own walls, meaning they have yet to find the right search tool for their business.

Given the slew of tools on the market, how can you find the appropriate one? Answering that question involves a dizzying array of factors, but three critical things to examine are:

- **Connectivity:** the ability to easily navigate around internal and external (Web) data from a variety of specific sources.
- **Effort:** the ability to implement, maintain, customize and use a solution with minimum expenditure of manpower and time.
- **Scalability:** the ability to accommodate large volumes of unstructured and structured data to keep pace as the company evolves.

These criteria offer a framework for assessing your company's data situation and how it may change over time, the ultimate concern when it comes to information access. These are also the criteria that most vendors highlight, so here are a few benchmarks for assessing your needs and vendor promises.

Connectivity

The amount of data that companies generate is staggering and shows no sign of slowing. IDC estimated that digital content and replicated data exceeded 281 exabytes in 2007 and expects it to grow ten-fold before 2011. Moreover, the type of data to be searched and indexed is changing from mostly document-based structured data to a combination of structured and file-based, unstructured data (such as rich media). In a 2007 study by the Taneja Group, 73% of users surveyed indicated that 60% or more of their data is unstructured.

What's more, the information that must be queried is all over: some on local devices, some with databases and legacy applications, and still more outside the firewall with software-as-a-service applications and on the Web.

As such, connectivity to different data sources has become a major factor in successful enterprise search implementations.

“The amount of data that companies generate is staggering and shows no sign of slowing.”

Paul Doscher, US CEO, Exalead

Answering a few basic but important questions will take you far in determining the level of connectivity you need:

“The information that must be queried is all over: some on local devices, some with databases and legacy applications, and still more outside the firewall with software-as-a-service applications on the Web.”

Paul Doscher
US CEO, Exalead

- What kind of information does your organization need to succeed?
- What context is required to give the information sufficient meaning?
- In what way should data be combined, integrated or processed?
- How many and what kind of users need to access this data?
- What will the access frequency be?
- Since connectivity has implications for performance and ease of use, what kind of indexing capabilities and results presentation does your business decision-making demand?
- How do you imagine your business—and your data pool—changing over time?

Solutions designed to meet the broad information access needs of larger corporations—thousands of users, hundreds of queries per second, documents in the high millions—offer essentially unlimited connectivity. Of course, the downside is often complex search interfaces that can impede intuitive search, and the platforms can be expensive to customize, integrate and deploy. The key is finding a solution that gives you the connectivity you'll need with the simplicity you want. This is where effort comes in.

Effort

The ability to connect to disparate data sources is one leg of the stool. The ability to easily manage and configure both system and data is the second. There are two types of effort worth considering: manpower required to implement, configure and maintain the tool, and the energy required by the user to use it. With regard to manpower, obviously the best technology investments are the ones that will perform their function well without becoming a drain on already busy staffers or already stretched resources.

Moreover, users accustomed to consumer Web search are increasingly unhappy with rigid, drill-down information retrieval tools embedded in many enterprise applications. They are equally unhappy with the meager, Web-style keyword search tools. In a business environment, sifting through an endless, flat list of links to find the information you need is unacceptable. What workers need is the simplicity of launching their quests with a keyword text box which returns organized, navigable results that can guide them to the information they need. When considering the effort your business is willing to support, consider these questions:

- How many people will be required to manage the tool, and will they be trained subject matter experts or can they be part-timers who multiplex across other duties?
- How stable is the solution and what day-to-day maintenance is required? Do you need a high degree of administrator intervention?
- How much control do you need to exercise over the management of your information?
- How often do changes need to be made to the way you access information?
- What sort of results presentation do you need to make the fastest business decisions?

“In a business environment, sifting through an endless, flat list of links to find the information you need is unacceptable.”

Paul Doscher

It is also important to determine what ongoing support will require. For example, if you want to add another content repository to a search application that is already running, do you really want to bring down the whole application and re-index all the information as if starting from scratch? Probably not, so check to see how your search solution works when you're in production. Which brings us to scalability...

Scalability

The third leg of the search stool is scalability. Interestingly, two major factors of scalability are its connectivity and the effort required to implement and maintain it. A scalable solution connects to a variety of data sources, requires limited touch to keep it functioning and can be easily configured and reconfigured as needed.

Here are a few questions to ask to make sure a solution can adapt to the business rather than vice versa:

- How much control do you have over applications? How extensible is the solution?
- Can you add search fields/data sources to the index without having to rebuild?
Can you modify any index field without having to re-index the entire document?
- Can you add hardware as needed without having to reconfigure your existing system?
- How complete is the API?
- How easy is backup and rollback?

These baseline questions will help you home in on the "right" search solution that will optimize your ability to access information crucial to business decision making. The bottom line: You want to find a tool that is easy to deploy, manage and use but that offers levels of scalability, performance, sophistication and connectivity that fit with your IT environment. This way, you can really make your data sing.

"A scalable solution connects to a variety of data sources, requires limited touch to keep it functioning and can be easily configured and reconfigured as needed."

Paul Doscher, US CEO, Exalead

About Exalead:

Founded in 2000 by search engine pioneers, Exalead is a global software provider in the enterprise and Web search markets. Exalead's worldwide client base includes leading companies such as PricewaterhouseCooper, ViaMichelin, GEFCO, American Greetings and Sanofi Aventis R&D, and more than 100 million unique users a month use Exalead's technology for search.

Today, Exalead is reshaping the digital content landscape with a platform that uses advanced semantic technologies to bring structure, meaning and accessibility to previously unused or under-utilized data in the disparate, heterogeneous enterprise information cloud. The system collects data from virtually any source, in any format, and transforms it into structured, pervasive, contextualized building blocks of business information that can be directly searched and queried, or used as the foundation for a new breed of lean, innovative information access applications.

This revolutionary technology and others make Exalead one of the most important companies to watch in the fields of digital content search, discovery, management, security, and storage.