

Interview with Stephen E. Arnold

Jess Bratcher, Bratcher & Associates

Contact: Stuart Schram IV, 502 228 1966

Interviewer's Note: Stephen E. Arnold is a consultant specializing in online information. His clients have included the U.S. government, the world's largest software company, and investment banks. He is the author of the new study "Beyond Search: What to Do When Your Enterprise Search System Won't Work," published by the Gilbane Group in Cambridge, Massachusetts. Arnold's two studies of Google – "The Google Legacy" (2005) and "Google Version 2.0" (2007) – changed the public's perception of Google from a Web search program to a comprehensive application platform. In 2003, he started writing the "Enterprise Search Report." He authored the first three editions of that seminal analysis of more than 30 vendors of enterprise search, a process often described as Intranet search or behind-the-firewall search. The Gilbane Group interviewed Arnold about his new study and one of the hottest enterprise software applications: search and retrieval.

Gilbane: Why "beyond search"?

Arnold: Four or five years ago, search was plugging key words into a search box. Today, in an organization, if my research data are correct, more than half of a system's users are dissatisfied with their search systems. Key word searching is frustrating, not a novelty. Today the volume of information in an organization is vast and getting larger. Most search systems don't deliver what users want.

Gilbane: What creates this information access problem?

Arnold: Vendors are delivering systems that are often very complicated. The company licensing these systems often doesn't understand the human, technical, and financial resources required to keep these beasts under control. The vendors are trying to deliver on-the-fly classification, entity extraction, and rich content processing. The licensees want these functions, but find these systems too expensive to operate. Most search systems are sub-optimal because of a lack of resources.

Gilbane: How widespread are search problems in organizations?

Arnold: My research suggests that about two-thirds of the users of an enterprise search system are dissatisfied with their incumbent system. The French business intelligence company Sinequa in Paris, France, released similar findings in April 2008. One of the descriptions of their data suggested that most organizations operate "information graveyards" because search does not meet the employees' and users' needs.

Gilbane: Who's responsible for the problem?

Arnold: Some of the people covering search for giant consultancies like Gartner, Forrester, Ovum, and others dance around the problems of search and information access. It's easy to blame technology and then say that more and better technology is needed. I don't think this is a technology problem. I think the vendors and the customers share responsibility for the deteriorating situation.

Gilbane: What are the vendors doing?

Arnold: The vendors are trying to make sales, generate revenue, and satisfy customers. In the effort to make a sale, vendors describe their systems and show compelling demonstrations. I can think of only a small number of vendors who take the time to educate their prospects and present their systems' strengths and systems in an objective way. With more than 150 vendors selling search and content processing, organizations have a greater chance of learning about search from vendors who focus on buzzwords and slick demonstrations, not the nitty-gritty details of their specific system. I can't fault vendors for making a sale, but I think some of the sales techniques plant seeds of dissatisfaction. By over-promising, licensees feel short changed when they discover that some features and functions cost a great deal to implement or involve work procedures that the licensee cannot perform.

Gilbane: And what about the licensees? How can a customer contribute to the poor performance of the enterprise search system?

Arnold: I think buyers of search systems are often under pressure to make a decision. As a result, it's easy for an IT manager or a procurement team to pick a system based on the vendor's name recognition, the personality and charm of the sales person, and marketing sizzle. When the system is installed, the licensee learns the true cost and complexity of the system. That's when the trouble begins.

Gilbane: What do you mean by “that's when the trouble begins”?

Arnold: Let me give you an example. Most organizations don't know how much content has to be indexed. Not only that, most organizations don't know how much of that content is changed or updated every day. When the search system is installed and the first indexing is completed, there's a great deal of learning about the content that takes place. Security issues previously hidden are front and center. Addressing these issues can take time and money.

Gilbane: Autonomy has more than 17,000 customers. Fast Search & Transfer has several thousand. Google has more than 9,000 Google Search Appliance customers. Are these systems not working?

Arnold: There's a difference between user satisfaction and a system indexing and serving results. Dissatisfaction means that the system doesn't do what some users expect, want, or need. When properly resourced, these vendors' systems work. The problem is that users are making it clear

that whatever system their organization uses for search, that system doesn't do what some employees require to do their jobs quickly. Lousy search, from the employee's point of view, is now a growing problem.

Gilbane: I'm still struggling with why search doesn't work for more than 60 percent of the users in your study and the same number in the Sinequa study.

Arnold: I think people forget how tough it is for an employee to formulate search queries that deliver exactly what he or she needs. The point-and-click interfaces help because people are good at recognizing something that might be relevant. But most point-and-click interfaces require quite a bit of tuning and tweaking. Even the best tools may have some drawbacks that can't be understood until the system is deployed and available to real users. Obviously, if two independent studies report that 60 percent of a system's users are dissatisfied, it's clear there's a problem.

Gilbane: What are some of the fixes available to a company with a search system that doesn't work?

Arnold: Let me hit a couple of key actions. The “Beyond Search” study is about 300 pages devoted to answering this question, so it's complicated, because search and content processing are among the most challenging functions in computer science to get right.

First, there are new, reasonably-priced systems that can be dropped alongside an incumbent system. Once the new system is up and running, the incumbent system can be moved to the sidelines or turned off. This is the “rip and replace” approach, and there are some vendors who offer products that can do what some of the better known vendors can deliver in a more modern, light-weight package.

Second, there are vendor who have utilities or add ons. These can operate on top of an existing system to make it better. If you follow this approach, it's true you are adding a third-party component, but in some cases, that's preferable to the “rip and replace” approach.

Third, you can live with your existing search system, but you will have to step back and decide if you need to turn off certain problem functions. These are ones that slow throughput or require more time and money to improve than your organization is willing to spend. This is a popular approach because it goes back to basics. A system that works predictably is better than one that times out or can't update its indexes.

Fourth, you can go to departmental or scaled-down systems that handle a specific content domain. Let me give you an example. The legal department in a big corporation needs to manage information related to a legal matter. Instead of trying to make the enterprise system do what the attorneys need, it's easier to license an eDiscovery system and let the lawyers have a specialized system that meets their needs.

Gilbane: Which of these approaches is a better and cheaper approach?

Arnold: There is no “one best way”. Each organization has different needs, different information flows, and different information access problems. In “Beyond Search” I profile 24 vendors' systems that can deliver a better search experience. I'm just not a fan of the search system that does everything from key word search to business intelligence. Too many vendors and licensees want to “boil the ocean.” I also favor the more modern systems developed since 2001 or 2002, not the systems from the mid-1990s or earlier.

Gilbane: You put the phrase “enterprise search” on the map with your “Enterprise Search Report.” Now you are saying “beyond search.” Were you wrong about enterprise search?

Arnold: No, times change. With more employees becoming search savvy, vendors and procurement teams have to look for ways to solve problems. We have to move “beyond search.” That's the focus going forward for the next two or three years. After that, I don't know what the challenges will be in search and content processing.

Gilbane: Who should read “Beyond Search”?

Arnold: Anyone interested in improving their existing search system, vendors, procurement teams, investors in search companies, and entrepreneurs.

###