

It's your power  
over information



## Problems of Corporate Search Solved

---

© SearchInform Technologies Inc.



### **Introduction**

Today when IT innovations continuously change the way we communicate, information has become a vital resource and those who can use and manage it are definitely ahead. Thus one of the major challenges facing companies at present is the need for quick information search in large data volumes. The organization of data access is in direct relation with the technologies and software that are quick and efficient in processing information. This section lists only few major problems solved by SearchInform technology. It should be noted that not one of other existing technologies provides an efficient solution to any of them.

### **Prompt Search of Required Information**

Conventional phrasal search implemented in all existing systems (both desktop and corporate solutions, as well as in the Internet) cope with their task, but the result is far from being satisfactory. One of the stumbling blocks is the time wasted on selecting key words and viewing useless documents displayed in the final lists of search results by a not quite correct query word (phrase). Search for documents with similar content implemented in software based on SoftInform Search Technology cuts down the time spent on information search to the minimum, providing accurate similarity results and displaying the list of documents in a convenient format.

In other words, we are not talking about a system that will be the fastest to react to the query (1 second instead of 5 seconds). We are saying that you will have to go over a much smaller amount of texts in order to find just the information you need.

Let's say you need to find documents (in particular, news) on a merger of two IT companies or prospective firms stored in your informational system or database. The conventional procedure looks as follows: you type in "company merger" and get a huge list of documents (though not necessarily exactly what you were looking for). Further on, you will find the document you were looking for in the 10th place. Having looked through this document you understand that it would be a good idea to try searching by phrases "company takeover" and "company acquisition", etc. As a result you end up thinking up more key phrases for searching and processing an enormous number of documents.

But when you use the SoftInform technology, as soon as you find the required document you can use the similar search feature by clicking on "find documents with a similar content", and you will immediately see a list of relevant documents (with a similar content) dedicated specifically to the required topic. Thus, instead of wasting several hours on searching (browsing through the list of results and trying new key phrases) you can do it all within a matter of a few minutes.

### **Fuzziness of Informational Content**

An enterprise's database or information system may contain documents from various sources, containing similar or identical information. The same text may appear under different headlines, with slight changes or amendments, which only make it more difficult to find what you need. For instance, a database may contain two or three similar documents with a virtually identical content, but with different titles and slight alterations in the text itself. And what about the situations when one expert will comment on document #1, while another expert comments on document #2, and so on. First of all, it means double work (why comment two or three times on the same document?). Second of all, if, say, the comments are different, part of the



process or some of added information may be lost or extremely hard to find. SoftInform Search Technology provides a perfect solution to this problem: when someone adds a new document to the database, by means of the similar search tool he can determine in an instant whether this document is new or whether it duplicates a file already existing in the database.

### **Generating Similarity Report**

It is not uncommon to find duplicating documents taken from different sources or added by different people in the enterprise's information database. As a rule, information is accumulated over years; therefore, in order to use the strengths of SearchInform technology to the full extent, you have to rid your system of unnecessary duplicates. Practice shows that those unit managers who had tried to rid themselves of unnecessary duplicates were stunned at the disorder that had been deeply set in their data.

In order to find duplicates and unnecessary "similar" files SoftInform Search Technology uses the feature of similarity analysis reports. The process itself is dozens of times less time-consuming than conventional comparison. For instance, comparing documents in an information database that contains, say, several million documents will take about a month. But with the similarity analysis report not more than a day.

### **Consolidating Information from Various Sources**

This problem deserves special attention. While the quantity of information experiences a rapid growth, large enterprises have to invest a lot of efforts into combining information from various systems into a single one. In addition to the high cost of developing new solutions, another major problem they normally get confronted with is deploying this solution at the enterprise. This, in its turn, may be quite painful for the management during a certain period upon deployment. SearchInform technology and information systems developed on its basis solve this problem absolutely painlessly for the enterprise.

SoftInform Search Technology is first and foremost a corporate tool for searching information in the enterprise's local network (documents of virtually any format), information systems, DBMS, CRM and so on. SearchInform has an irreplaceable ability to consolidate information from various sources. SoftInform Search Technology incorporates a rubricating system (a fast and convenient tool for categorizing any documents by required topics) and automatic categorization of documents, i.e. automatic distribution of new documents into existing rubrics according to the general similarity principle. Such an approach to organizing and consolidating data lets you structure the information components of any large-scale enterprise within the framework of one application. On top of that converting documents and data into one single format won't be necessary. All information available for indexing and further search can be categorized, structured and conveniently displayed.

Due to such an approach SearchInform technology is easily integrated into an already functioning information systems of the enterprise, which helps it solve all problems connected to consolidating and searching data from various subsystems without having to re-organize the whole informational infrastructure of the enterprise. Deploying SearchInform won't require changing the existing business processes and will allow the company to save the efforts it has invested into the already existing informational infrastructure. At the same time the system combines isolated corporate applications and data into a single information system, thus resulting in a more efficient solution to business tasks.

It's your power  
over information



### ***All in all***

To sum it all up, SearchInform is a technology designed for quick and accurate information search in large data volume. Its innovative and problem-solving feature of similar text search makes it stand out among all the other similar products present in the market. High indexing speed (15 to 30 Gb/hour), index size of 15-25% of the actual text information size, support of over 60 widely used text formats, Outlook and other email clients, mp3 and avi tags, as well as Microsoft Messenger and ICQ 99-2005 logs, indexing of all databases — all that and a lot more enable SearchInform to solve problems of both corporate and individual users.

***SearchInform is your ability to act while others are still searching.***