INFOLUTION
SOLVING INFORMATION COMPLEXITY

Archiving

The amounts of data that organisations need to manage grow exponentially. It has been estimated that an average company will double the amount of information it must manage every six to twelve months. Some reasons for this growth include external legal obligations requiring long-term data storage. Internal growth also requires companies to manage more information about their products, suppliers, partners, customers, employees, etc. Companies are confronted with data growth in both digital and non-electronic forms. In order to increase productivity non-electronic information is transferred into digital data. Once data is digital it must be classified, stored and retrieved effectively and efficiently.

This data growth challenges companies to come up with intelligent archiving solutions. Unfortunately, traditional solutions are static. They are based on fixed and hierarchical structures. Document handlers need to categorise documents manually. This is not only slow and cost intensive, but often causes errors in the process. In addition, retrieving information requires knowing where it was stored and this becomes cumbersome. Categories need to be defined upfront and can rarely be changed at a later stage. Also, it takes a significant effort to introduce new categories into traditional solutions. Overall, traditional archiving is inflexible. It keeps effectiveness and efficiency of the organisation at a low level.

Infolution Archiving is helping companies to become more effective and more efficient by offering an innovative architecture.
Infolution Archiving is flexible and easy. It categorises documents automatically at a high level of precision. Furthermore, there is no additional effort required to re-classify any document at any point in time in the future. This flexibility allows an organisation to be agile. It is easy because the document handlers do not need to be experts, the solution has internalised and automated the expertise, which otherwise would reside within the brains of the archiving people. Once the documents are stored, the **Infolution Archiving** solution enables inexperienced users to retrieve data accurately and with high precision. The solution can safeguard the data by determining who gets access to the various levels of data, since compliance to company policies, as well as regulatory policies are critical. Today's investment is protected for the future since **Infolution Archiving** adapts to changes at any future point in time.

The solution is applicable to structured and unstructured text data. It provides unlimited scalability, is built on industry standard platforms, and offers multi-language support. The open API allows OEMs of archiving solutions to easily integrate **Infolution Archiving** into their existing solution thereby significantly enhancing their capabilities to automatically classify their data. The retrieval function uses semantic search technologies.

The Infolution approach turns out to be very flexible and adapts easily to real life problems. It actually opens the door to answer questions that traditional archiving techniques would not be able to answer, unless the exact answer is given in one document.

The Infolution software suit is perfectly positioned to address today's archiving challenges.
The innovative solution is unique in many ways. **Infolution Archiving** is represented by an SQL compatible database that is based on a semantic concept. The underlying idea of Infolution's storage solution is a structure that is not based on strict hierarchies, but on a behaviour that is similar to that of humans. Humans learn about new concepts by creating a relation to already known concepts. The Infolution software does the same. It uses a number of basic language rules to determine a relation between two concepts. In addition, a statistical analysis gives an answer to the question, how close two concepts are.

The Infolution software not only recognises content, it also interprets it. The resulting benefits for the user are impressive:

- Speedy access to information.
- Scalability with data sizes.
- "Filters" separate relevant from irrelevant data.
- Automatic categorisation.
- High accurary and relevance.

The Infolution system is an ideal solution for any kind of archiving task and sharing knowledge across a larger community. Semantic database, document classification and synthesising are powerful tools to allow fast and focused retrieval. Filters separating relevant from irrelevant data significantly increase the effectiveness of users.

Examples

HowardsHome is a company specialised in collecting information for their clients. They continuously tap into about 40,000 news sources and extract information per their clients' criteria. They have to archive hundreds of Gigabytes in a way that any information can be retrieved again instantaneously, without enforcing multiple searches and without distracting the user with irrelevant data. The user should stay focused on his tasks, rather than fighting technology.
HowardsHome has selected Infolution software because of its scalability, speed, flexibility and its openness.

Infolution's partner, **BCT**, has customers who are facing different challenges. One of these challenges involves the need for an automated solution that classifies and archives incoming mail. The software should "understand" a letter, direct it to the right person within the organisation and even monitor the process until completion. **BCT** integrates Infolution software into its Automated Document Handling Solution, in order to help public authorities, as well as corporations, like energy provider Essent, to manage their mail flow. The number of documents being processed by this solution (structured forms or unstructured letters) exceeds hundreds of thousands each year. This archiving solution plays a crucial role in the process, which enables Essent to reduce the number of document handlers, as well as reduce the error rate in classifying their documents. Essent has internalised the expertise, which previously was inside people's brains by using **BCT**'s Automated Document Handling Solution.

The technology

Infolution Archiving sits on top of the **Semantic Search Engine** which uses semantic algorithms, statistical rules (Boolean and Bayesian), linguistic technologies and self teaching and learning capabilities. There are four major components to it:

Harvest: Data collection from various sources is done with intelligent software utilising a semantic analysis of the text surrounding the actual link. This allows Infolution software to determine the context in which the data stand and ultimately determine the priority of the document for the user.

Analysis: The core of this module is the first ever scalable and commercially viable semantic processor. It operates similar to the human brain understanding and classifying content – and it is capable of learning as it goes along. In other words: not only does it understand the meaning of a user query and the context of a found document, but it can also correlate the two.

Storage and retrieval: An SQL compatible database is the core of the storage module, which is operating on a semantic basis. It allows retrieving more from a document that just the actual content. It allows, for example, summarising a document tailoring it to the intention of the initial request. This is possible because not only facts are stored, but also the relation between them.

Access Control: A conditional module is incorporated that will follow the access control list of the underlying data storage. Infolution uses this information to safeguard that both documents and indexes are only visible to authorised users.

The Infolution engine works continuously, conducting the above four functions in the background, and updating its files as soon as documents are added or modified. Its architected is such that it provides unlimited scalability and high performance which are critical for searches in large websites or across content rich multiple websites. The engine prepares data for the main functions of the Internet search solution: search and summarize.

Document classification: Infolution's software automatically classifies all documents for storage and easy retrieval. This dynamic process allows changing or adding categories at any point in time. Categories are added dynamically to the database. There is no additional effort required to re-classify documents.

Document classification has a tremendous advantage for users. If a user searches for a term with multiple meanings, the classification scheme would allow him instantaneously to narrow the search down to the meaning he is interested in.
A search for "jaguar" would result in finding documents on the animal, as well as, documents on the car. Limiting himself to only a single category would help the user to exclude unwanted and meaningless data from his search.

Document summary: During the retrieval process, users will encounter one of the most visible benefits of the system: its capability to automatically create a summary. This feature is one of the most spectacular elements of the Infolution software. It is based on a document synthesizer which can take any number of documents and present a summary as one virtual document. The summary can easily be changed in length without loss of coherence. Also, it can easily be rewritten at any point in time placing an emphasis on different topics as required by the user.

The following features of **Infolution Archiving** make it a very powerful solution for companies of any size:

Scalability: Infolution supports up to 8 WWW servers, 128 file locations, 4 CPUs and allows more than 1000 users to share knowledge.

Industry standard platforms: Allows integrating Infolution technology into existing infrastructures. Not only can it be installed on existing servers, but also does it come with a documented open Application Programming Interface (API) to allow for tighter integration into existing software solutions.

Multi-language support: Allows international companies to standardise across borders. In total 72 languages are supported by Infolution.

User Interface: The rich user interface is especially designed for professional use. It supports the user with dynamic context interactive visualisation. The search engine finds the conceptual context of a search and gives the user a number of intelligent alternatives by relating queries. The contexts of the results are shown in an intuitive three-dimensional graph, in which more important concepts appear to be closer to the user than less important ones. Using visual cues, like mouse-over effects and animation, the system clarifies how and why certain concepts appear.

Infolution bv
Schouwschuit 37
1613 CJ Grootebroek
The Netherlands

Phone: +31 (0)228 52 40 70
FAX: +31 (0)228 52 40 74
www.infolution.com

INFOLUTION
SOLVING INFORMATION COMPLEXITY