



Unternehmen Leben



DAK migrates their information archive to PDF/A with LuraTech

The DAK, Germany's second largest healthcare company, has migrated their internal information archive to the new ISO standard PDF/A with help from LuraTech. The archive used to exist as digitalized microfilms, scanned into TIFF format. This had the disadvantages of not only requiring a lot of memory space, but in particular the files could not be full-text searched. Now, after converting and compressing approximately 300,000 pages of text, users of the DAK Info Service have direct access to all saved information. In addition, the DAK gained early knowledge with the PDF/A format, that they will put to use in future projects.



The DAK, as Germany's second largest healthcare provider, has some 4.7 million registered members and insures around 6.2 million people under the motto "Life Company". Approximately 12'000 employees across the country work for the company. With over 230 years worth of tradition and knowledge in the area of social insurance, the DAK helps in more than just cases of illness. The company is seen as particularly innovative within the regulated

healthcare industry. They reward good health through offers like a bonus program, integration contracts and health-related travel. Their standing is confirmed in the current special edition of the consumer magazine "ÖKO-Text", which ranks the DAK number one among health insurance companies.

The answer to all questions: the DAK INFO-Service

The internal INFO-Service of the DAK, located in the Hamburg headquarters, is the information platform for the employees in all branch offices. Inputs from all relevant professional publications flow into it. These are typically legal or medical articles, for example articles concerning legal aspects of providing healthcare, patient rights, malpractice cases, new treatment methods etc. The information is entered in the DAK literature database DOR!Sonline (DOcumentation Retrieval und Information System) where it is then made available to the internal INFO-Service for inquiries. Employees of the INFO-Service answer queries from the branch offices that are of a legal or medical nature, and place the information they process in the archive. "In addition, we have a duty to inform our employees about new laws and other develop-



ments that may have an impact on their work" explained Hans-Hinrich Harms, documentation officer of the INFO-Service for DAK. "This information is also managed in our database." A large number of the relevant articles, which sum up to about 3,000 per year, have 50 or more pages. These are digitalized using an external scanning service provider, "Alphacom" in Hamburg, and are sent as digital files back to the DAK. Smaller articles are scanned in-house using Hewlett Packard scanners and then filed.

Full-text search desired

Up to now, the scanned information was archived in TIFF format. Also to be found in the archive were TIFF files created from microfilm from the years 1988 to 1995. Afterwards, these files were stored in PDF format, which however offered no standard format for long-term archiving. A major problem was that the TIFF files could not be full-text searched, making it extremely difficult to search through them. For this reason, a decision was taken to look for a modern format and a suitable solution which the archive could be converted to. At the CeBIT in 2006, the DAK learned about the new ISO standardized PDF/A format for long-term archiving. Unlike TIFF, PDF/A supports text, reproduces colours like the original, and has a smaller file size. Compared to the native formats created by desktop applications like office or CAD systems, the reproducibility of PDF/A is guaranteed in the long run and requires only one viewer - the Adobe Reader.

The DAK found what they were looking for during the CeBIT at the booth of LuraTech, a company from Berlin. The compression specialists had just finished upgrading their product portfolio to support PDF/A. Scanned and digitally created documents can be converted to PDF/A using LuraTechs solutions, and the conformity can be verified. The latter is necessary to ensure that documents destined for the archive fulfill all requirements of the standard. The PDF/A syntax, a more restrictive variation of PDF 1.4, defines in detail what content is permitted and what is not. For example, the entire content must be contained in the file. External dependencies and references to external sources, which can eventually hinder long-term reproducibility, are forbidden. The fonts that are used in the document must be embedded in the file, and encryption is not permitted.

Poor legibility not a problem for the PDF Compressor

Using their LuraTech PDF Compressor Version 4, which converts scanned documents into highly compressed and full-text searchable (OCR) PDF/A files, LuraTech migrated the entire archive to PDF/A in just a couple of weeks. The files on microfilm were difficult to read and led to respectively poor quality TIFFs. Despite this, a very high character recognition could be achieved thanks to the advanced OCR technology that the PDF Compressor uses. In order to facilitate the uncomplicated classification of the converted files, LuraTech retained the existing file names. These consisted of the year that the file was created and a 5 digit sequential number. Only the extension (.TIF) was changed.

Satisfaction on all sides

Employees of the DAK INFO-Service profit immensely from the quick full-text search feature. It is also planned to eventually make all the information in the DOR!S Database available to the employees over a portal. The first steps have already been taken in this direction. At the same time, the project

was a type of test for the freshly introduced PDF/A format, for whose possibilities the DAK is very interested.

Hans-Hinrich Harms summarizes: "PDF/A provides a modern format, that offers more features and is also an ISO standard. The full-text search capability was the greatest advantage for our INFO-Service. The cooperation and communication with LuraTech proved to be very pleasurable. In addition, we discovered that colour scanning is supported today with PDF/A and we are examining the possibility of also providing colour documents in our INFO-Service in the future."

The ISO 19005 standard (PDF/A) defines requirements for creating documents suitable for archiving, based on the widely available PDF format. The standard specifies in detail what type of content is allowed and what is not. This and other specifications are intended to ensure the long-term readability of documents regardless of the application software and operating system in which they were originally produced. Experts predict that the designation of PDF/A as an ISO standard will promote the popularity of the format, and advantages such as full text search capabilities will lead to it replacing TIFF in the near future.

LuraTech Europe GmbH

Kantstr. 21

10623 Berlin

Telephone: +49 30 394050-0

Telefax: +49 30 394050-99

info@luratech.com

<http://www.luratech.com>