ABBYY®

ABBYY Recognition Server Technology Provides Quick Access to Official Archive of Documents

> *"ABBYY Recognition Server has been easy to integrate into our software/hardware complex, reducing the overall cost of the project and its implementation time. The results which have been obtained with the software meet the high standards of quality set by Xerox when providing services to our clients."*
>
> *Olga Efimova, PM Sector Manager*
> *Industry Solutions & System Integration Center, Xerox*

ABBYY Recognition Server has been successfully used to convert an archive of paper documents in the key office of state into electronic form.

Many federal government bodies have amassed huge archives of paper documents which are still relevant to today's decision-making processes. Due to its technologies and vast industry experience, Xerox has been chosen to carry out a large-scale archiving project to convert an archive of a governmental organization into electronic form.

Previously, whenever federal officials needed a document from the archive, they had to request an archivist to retrieve it based on the document descriptions and the archive inventory, all of which were available on paper only. Once the document was found, the archivist made a copy, which was then sent to the requesting official. Officials had to wait one or two days before they had a copy of a document in question from the archive.

To speed up the process, an electronic archive was commissioned. The project was carried out by Xerox Corporation using its own set of software and hardware tools coupled with ABBYY Recognition Server.

As the first step, all the archived documents which totaled over half a million pages — were scanned, producing electronic copies, or images, of 3,800 cases. Then, all documents had to be indexed in accordance with a predetermined data pattern. An important thing was to create an annotation for each document, i.e. a brief summary of its contents which had to conform to certain rules previously agreed by the Customer. To eliminate manual printing work and to make a full-text search possible, the electronic images had to be processed using an Optical Character Recognition system which extracted text and created searchable PDF files.

ABBYY® Recognition Server 2.0

www.ABBYY.com

The crucial OCR stage was performed by ABBYY Recognition Server, a high-performance server solution which automates OCR of images and PDF files. Xerox used the default Recognition Server package, which has the standard set of features, is very easy to use, and does not require any assistance from ABBYY specialists to deploy and set up.

The documents were automatically recognized 24 hours a day, completely unattended. Each page took 15 seconds at the most to be processed. Then, Xerox operators selected and copied the most important text data from the processed documents to create annotations.

The use of ABBYY Recognition Server at annotations assignment to archival documents within archive digitization process speeded up the entire conversion project by 25% and reduced its costs by 7.5%.

As a result, 3800 cases containing over 570,000 pages were scanned, recognized and supplied with annotations. The resulting electronic data were exported to the huge corporate electronic archive.

Now it takes the organization's officials only minutes to find and retrieve the documents they need, instead of days in the case of the traditional paper archive. If an official requires a document from the archive, all they need to do now is to open the electronic archive access window on the screen of the computer and type in a query based on the document's attributes or annotation (or specify key words for a full-text search). Within seconds, a list of matching records appears on the screen, and a user can open and print the necessary document out.