

## History now available digitally

As a researcher, archivist, historian or librarian, you are a guardian of historical documents. You understand the challenge of preserving such documents in scarce space while making their content accessible. In pursuit of that goal, massive scanning and data capture projects have been initiated and are still ongoing



### **'Search-engine-like' display of information found in complex historical documents**

Scanning books and archives is not the end of the preservation process; it is the beginning. Understanding the content and making it digitally searchable as information rather than just providing digital images for manual review is the end of a successful conversion process.

That is where ArgusSearch can help. ArgusSearch allows 'search-engine-like' display of information found in historical and complex handwritten or hand-printed documents. Based on Planet's leading proprietary ICR technology, ArgusSearch enables information providers such as genealogy search sites, archives and libraries to make the content of their historical treasures digitally available for research. Unlike conventional ICR that provides either a transcribed word or a 'not read' result, ArgusSearch provides 'quasi' matches for a particular search term based on the equivalence of the search term and a particular instance of the term within a document. This allows a much deeper level of penetration of the information locked in a historical document and enables historians, researchers, and the general public to search for relevant information in archives or libraries in a 'search-engine-like' fashion.

ArgusSearch is ideal for large volumes of documents of cursive handwritten, or historical handprint (constrained and unconstrained). It processes any input document type, including obituaries, newspaper articles, census data, or passenger manifests. At the core of ArgusSearch lies the ICR industry's most advanced recognition and classification software, which is able to process even the toughest and most complex input documents. Typically, OCR/ICR scans documents and converts their content into a machine-readable format (ASCII). This approach has limitations when it comes to historical documents written in complex script or by multiple writers because the text does not always allow for correct deciphering and transcription of individual words. ArgusSearch solves this dilemma. ArgusSearch is a document content capturing solution for large volumes of complex, handwritten documents. The output of ArgusSearch is not a transcribed ASCII text, but a matrix of word characteristic features that can be sorted for their equivalence to a certain search term.

## Features

### Word spotting in historical documents

Given the large amount of variability in the handwriting data fields and the low quality of captured images of the material, historical documents are traditionally transcribed by hand to make them digitally available, especially by search. Each word must be manually deciphered and annotated. This process is time consuming, costly, error prone and tedious. ArgusSearch's word spotting capability greatly reduces the amount of work required, speeds up the process, and makes large volumes of historical documents available for digital search and retrieval.

### ArgusSearch contains the following features:

- Available as a stand-alone solution, or can be made available as a web-based service
- Interfaces with multiple input platforms, such as scanners, cameras, or mobile devices
- Searches within unstructured or historical documents
- Facilitates configuration and optimization of search rules
- Scales up to meet large volumes of content

### It provides applications for:

- Government Archives (Manuscripts)
- Libraries
- Genealogy Research (Census Data, Obituaries, Passenger Manifests)
- Universities
- Museums
- Data Service Providers

## Planet IS Inc

At the cutting edge of neuro-informatics, Planet Intelligent Systems develops biologically-inspired computing technologies featuring human-like intelligence. Planet's system architecture is constantly evolving and always integrating the latest concepts of neural network computing and artificial intelligence to push our propriety and powerful algorithms at the center of our solutions to new performance levels. Planet's writing recognition technology is at the leading edge of technology, winning prestigious prizes at the 2014 International Congress on Frontiers in Handwriting Recognition (ICFHR) as well as the 2013 International Conference on Document Analysis and Recognition. The unrivalled capability and accuracy of our data capturing solutions in postal automation, traffic surveillance analysis and document conversion are a testament to the breakthrough capacity of our technology.

### PLANET intelligent systems GmbH

Residence Park 1-7  
D- 19065 Raben Steinfeld  
fon +49 (0)3860 5010-0  
fax +49 (0)3860 5010-250  
email info@planet.de

### Planet Intelligent Systems Inc.

1000 Geneva Ct.  
El Dorado Hills, CA 95762  
ph: 916 941 0930  
fx: 916 673 6531  
email: info@planet.de

