

Bringing history to life

As a researcher, archivist, historian or librarian, you are a guardian of history. You understand that converting valuable historical records to digital data is critical to preserving their integrity, to managing limited space, and to bringing their history to life.



Historical records



Historical records

Ability to process even the toughest and most complex input documents

Scanning books and archives is not the end of the preservation process; it is the beginning. Indexing and classifying those records so that their content is searchable and widely accessible is a critical step to the success of the conversion process and to their relevance in the future.

Yet, manually capturing information from historical documents is very costly, error prone and inefficient. That is where ArgusDiscovery can help. ArgusDiscovery locates, recognizes and indexes historical documents, even those with cursive handwriting or historical handprint (constrained and unconstrained) handprint, scripts and fonts. ArgusDiscovery processes input document types such as obituaries, newspaper articles, census data, passenger manifests, and other freeform documents. At the core of ArgusDiscovery lies the OCR (Optical Character Recognition) industry's most advanced recognition and classification software, which is able to process even the toughest and most complex input documents.

Until now, historical documents are typically transcribed by hand to make them digitally available. This is due to a great amount of variability in the handwriting and the low quality of images captured of the material. Each word has to be manually deciphered and annotated. This process is time consuming, costly, error prone and tedious. A software application for word spotting in historical documents can greatly reduce the amount of work required; it can also speed up the process and make large volumes of historical documents available for digital search and retrieval. At the core of ArgusDiscovery's word spotting and indexing capability lies a powerful ICR (Intelligent Character Recognition) engine that is specifically designed to locate, capture and index a pre-defined vocabulary.

Features

ArgusExtract interfaces with any input platform (scanners, cameras, mobile devices, digital images) and recognizes:

- Machine print (any font)
- All forms of handwriting, including constrained and unconstrained handprint and cursive handwriting (ICR)
- Historical fonts such as 'Fraktur', 'Gothic', and 'Sütterlin'

Sample projects and sources include:

- Data Extraction for repositories and online search applications
- Indexing, Sortation and Categorization of large historical volumes
- Data validation and verification
- Content generation for online data service providers

ArgusExtract provides solutions for the following industries:

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

The image shows a page from the 12th Census of the United States, Schedule No. 1 - Population. The document is a grid with multiple columns for demographic data. Handwritten entries in cursive script are visible throughout the grid. Three specific entries are highlighted with white boxes and black text: 'Peter Johnson', 'Porter John', and 'Porter John'. The document is dated 1900 and includes various administrative markings and headers.

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

