

Business Process Automation Solutions Document Classification

ancora was developed to bring simplicity to a technology that historically has been complex, over-priced and difficult to configure, install, use and maintain.

Do you need to scan mixed batches of structured, unstructured and semi-structured documents within a single batch and separate them without the need for separator sheets, barcodes or patch codes and extract critical data elements from these documents based on doc class?

With ancora and our patented classification engine you can scan, classify, index, and deliver data from nearly any form of paper or electronic document without the need for manual sorting or preparation

Accounts Payable Automation — automatically sort and separate PO based invoices from non-PO based invoices, utility invoices, credit memos and invoices by country of origin without any manual intervention, separator sheets or barcodes. Also, automatically identify supporting documents keeping them with their original invoice.

Human Resources, Mortgage Loan Documents, Medical Records — to name just a few applications our classification engine can automatically classify the different document types contained within these applications and apply data extraction rules to these documents based on the specific document type.

How does it work?

Using artificial intelligence and cognitive computing we train the system to classify the different document types. It is a mixture of computer science and cognitive science, the understanding of the human brain and how it works. By means of self-teaching algorithms that use data mining, visual recognition, and natural language processing (the understanding of common sentences while being able to sense the mood) the system is able to solve problems and optimize the human processes. Training is accomplished by simply dropping like documents into folders and applying the algorithm to learn what these like documents are. It uses content, layout, color, font sizes, etc., just like a human would use to identify like documents.

