
What is an OCR Program

What is an OCR Program

What is an OCR Program?

An optical character recognition (OCR) program incorporates several technologies and is used to read text from scanned images of documents. It usually can recognize a variety of text either handwritten or printed font, and in some cases can be used to recognize bar codes and other logos. Therefore, while generally OCR programs are used to recognize printed text and convert scanned images into machine encoded text they can do much more based on the technologies they have.

What Technology is Present in an OCR Program?

Several technologies are used in OCR programs and they allow OCR programs to do much more than just recognize printed font. When advanced technologies such as intelligent character recognition (ICR) and optical mark reading (OMR) are incorporated into an OCR program, it allows the program to read handwritten font and barcode. Most high end OCR programs in fact, incorporate ICR, OCR and OMR and have high recognition rates, as they have a self learning database akin to a neural network which is constantly updated with new recognition patterns.

Why Use an OCR Program?

Scanned images of documents are all well and good, and they can be input into document management systems, but several problems arise when these

scanned images have to be accessed. Scanned images are basically photos of documents and remain so even if they are converted into PDF files. They cannot be tagged or indexed and consequently cannot be searched for specific content. Users have to rely on document names when searching for files and this can result in important content being overlooked or remaining undiscovered. Scanned documents also cannot be searched, so users have to read through entire documents to find specific data, and once they find it retype the data into a word document. An OCR program can effectively and accurately deal with problems described above, and make data accessible to users faster than any human data entry operator. Output from an OCR program is usually in the form of searchable text which can be edited. Some advanced OCR programs even produce formatted files complete with logos and columns closely resembling the format of original scanned image.