

---

## Invoice File Processing

The process by which OCR extracts data from an invoice is extremely elegant. Initially, a blank, standardized invoice is fed into the software. When a filled up invoice comes up, the file is scanned and then compared to the image of the blank invoice. An invoice contains many fields where data is entered, and these fields are compared to see if they contain data. This data is then extracted from the invoice, and read using sophisticated OCR technology that compares the printed or handwritten data to stored bitmaps of text and numbers and creates strings of numerical or alphabetical data that can now be manipulated using any authoring software.

How is this data used?

Invoice processing requires a lot of mathematical calculations that is done on an input data file. Extracting this data from the invoice using OCR can make the calculations easier by enabling the use of electronic calculators that can be programmed to automate themselves with the extracted data. One can create spreadsheets of all invoice amounts for easy viewing, create graphs to control and calculate overall expenditure etc, and use the data in many other ways to help manage the finances of the company.

How does file extraction benefit invoice processing?

Data extraction from files can greatly benefit invoice processing. Firstly, it reduces the potential for errors by minimizing the human factor. Secondly, it saves on the labor and also on the quality of labor; very low-grade labor can still do automatic invoice processing, and do not need to be accounting degree holders. Thirdly, it reduces the time that is taken to process invoices, because a mechanical system can usually work much faster than humans. Fourthly, and as a cumulative benefit from all the above, digitization can reduce the cost accruing to a company by a significant amount.