

Factors Affecting Data Entry Cost for “Standard” Data Entry

In business since 1996, we learned very quickly every job is different and customer requirements vary widely. There is no such thing as “one-size-fits-all” and no quick answer to what it will cost to work on your project.

This article deals strictly with “standard” data entry -- work not requiring such services as data verification, data validation, Internet research, or data collection by reference to outside sources. It pertains strictly to entering data from client-supplied source documents.

Cost is a very important consideration in your selection of a data entry company, and the more information you can provide us the more reliable will be the quote you receive.

The methods we employ depend on the kind of documents provided and such factors as whether the information to be extracted is handwritten or typed. For handwritten data, we use live data entry operators to key the data from scanned images. For typewritten data, we may employ OCR (Optical Character Recognition), typically followed by some level of human verification, or “text layer” capture from fully digital pdf files.

With constantly improving capabilities offered by the Internet, the data entry processes have changed dramatically over the last several decades. In the early days of data entry, brick and mortar operations were predominant, with facilities considerations not only being a significant cost, but also a factor limiting scalability. Now, the Internet offers tremendous opportunities for remote, secure processing, faster turnaround times, and greatly enhanced scalability with limited incremental cost.

First, there was key-from-paper, then key-from-image, and then various automated recognition technologies (Optical Character Recognition – OCR; Intelligent Character Recognition – ICR for handwritten information; Mark Sense for checkboxes; barcode recognition, and others). All of them offer different processes, staff requirements, hardware and software requirements, as well as different cost and quality profiles.

Also, there was a major trend in using offshore companies to reduce costs. In 2013, we decided to work exclusively with U.S. staff, providing added security, more efficient processing, and better quality. This, however, does come at a higher cost.

In most cases, we charge a flat fee per data record or document. After project set-up, each invoice is determined by multiplying the number of data records or documents processed times the agreed flat fee per record.

Factors Affecting Data Entry Cost for “Standard” Data Entry

Axion takes a cost-based approach to develop a fee for data entry services. We typically do not charge based on hourly rates. Instead, we usually charge a flat fee per data record.

Once we arrive at an estimate of the cost per record or document, we then add something for overhead and profit. For the majority of the projects we work on, our fees range from \$0.25 (25 cents) to \$2.00 per data record; however, fees have been as high as \$22 per record.

A “record” is generally comprised of all the fields identified by the client as representing a complete, identifiable (and billable) string of information. In some cases, all the information on each source document represents a “record”(e.g. enrollment forms), and in others each entry in a listing can constitute a record.

The balance of this discussion will attempt to hit the high points affecting cost, and is divided into two general categories: 1) Traditional – work performed on our network using our system; and 2) Work performed on the customer’s system.

Traditional – Work performed on our network and system

All “standard” data entry work is performed using scanned images, either those provided by the client, or by our scanning partner. It is “heads-down” data entry where the task is literally “key what you see” with no other operations being performed.

When doing the work on our system, here are the basic processes and cost drivers:

Document handling: Most customers now send us scanned images due to cost, document security, and turnaround time considerations; however, customers not having imaging capabilities or unable to outsource the imaging, will send us hard copy documents. Factors affecting the cost include:

- Are documents being mailed one at a time, for example in connection with a direct marketing campaign, or in large batches?
- How often will we receive documents? Daily? Weekly, Monthly? All at one time?
- Will a P.O. Box need to be set up?

Document Scanning:

- Will any document preparation be needed, such as: opening envelopes, unfolding, removal of attachments, tear-off portions, staples, Post-It notes, stickers, etc?
- Document size - This affects throughput. Also, too small or too big cannot be scanned.
- Document uniformity - Are they all the same size? Are there differences in shading?
- Paper type - is the paper real thin or have a sticker attached?
- Document condition -- Poor quality documents will require extra time.
- Document arrangement - To be scanned efficiently, all documents must be right side up and facing the same direction. There can be no staples, paper clips, etc., and the

Factors Affecting Data Entry Cost for “Standard” Data Entry

documents must all be flat (unfolded). The cost of document preparation usually exceeds the actual cost of scanning.

- Batch integrity - Can we scan continuously, or do we have to retain client batch integrity?

Programming:

Once documents have been scanned, the images are imported into our data entry system. For each project, a "template" is developed within Axion's data entry system, programmed for the specific fields on the form. Each field can be programmed to restrict data entry to acceptable values. Any business rules a client defines for specific fields or forms can be programmed into each template.

One example of specialized programming we incorporate involves address correction and verification. Other examples include embedding formula validations or databased look-ups in fields. The factors affecting the cost of programming include the following:

- a. Document complexity --size, number of pages, number of fields per page, number of records per page
- b. Document uniformity -- are there multiple versions of the source document? Validations required -- for example, integrating client-provided lookup tables, address correction, etc.
- c. Format of output file required
- d. Programming is a one-time cost (unless the source document changes) that will generally average 500-750 for a single-page document.

Data Entry – Key-From-Image/OCR: Using the data entry template developed for a specific project, our data entry operators key the required data from scanned images. All work is performed on Axion's cloud-based network residing on the Google Cloud Platform (GCP).

Data entry operator workstations connected to Axion's cloud-based network function merely as “dumb terminals”. They cannot download or print anything. Using a split-screen, the operators view "snippets" of the scanned image in one part of the screen, and the data entry fields in the other part of the screen. The factors affecting the amount of time it will take to process a document, and hence the cost include the following:

- e. **Number of fields to be keyed**
- f. **Number of characters in each field**
- g. **Type of data** - numeric, alpha, alphanumeric
- h. **Complexity** – having numerous business rules that must be followed can slow things down, especially at the beginning of a project.
- i. **Form Design** - this is often overlooked but can have a huge impact on data entry **Legibility** - handwritten vs. typed; also, using a "constrained" form with each handwritten character printed in a separate box improves legibility; also,

Factors Affecting Data Entry Cost for “Standard” Data Entry

the quality of the source document may affect the quality of the scan and hence legibility.

All of the above affect the average operator production rate per hour. To the extent the customer has been doing the work in-house, we will use their historical average as a good-faith estimate upon which to determine the per record/document operator cost. The cost of the actual data entry is the hardest one to estimate, and most vendors will be reluctant to provide a firm fee quote without seeing, and if possible, testing a representative sample of actual source documents.

2. **Required accuracy** - For single-key data entry, our US workers generally achieve accuracy ranging from 96-98%, depending on the type of data (see above). If greater levels of accuracy are required, the client may opt for having us double key the entire document or selected fields. See <http://www.axiondata.com/quality.htm> for a full discussion of Re-key verification and other methods used to control accuracy.

3. **Start-up Costs:** A project manager is assigned to every customer, who will gain a full understanding of the project, business rules, staffing and turnaround requirements. The project manager will develop training materials, arrange for a training set of documents if possible, train the data entry staff and monitor performance. Depending on the complexity of the project and how aggressive the ramp-up schedule or completion requirements are, a separate fee may be charged to defray the associated cost. This is especially true in the case of one-time projects. For ongoing projects, long project duration and high volumes may mitigate the need for recouping this cost right away.

4. **Ongoing Project Management:** The project complexity, turnaround requirements, number of staff required, the incidence of “exceptions to the rule” that must be discussed with the customer, and the required frequency of transmitting data output files all affect manager time, the cost of which can be disproportionate to the actual data entry cost, especially for smaller projects.

Because every project is unique, and each client has different goals and objectives, it is not practical to take a 'cookie-cutter' approach to data entry pricing. The only good measure is to let us take a look at your project. That way we can give you a reasonable fee estimate based on the factors discussed above.

Factors Affecting Data Entry Cost for “Standard” Data Entry

The New Traditional – Entering Directly Into Your System

Increasingly, customers are asking us to enter information directly into their own system, typically a web-based program. In this way, the customer retains security over the documents and data, accelerates turnaround time, and eliminates the cost of having us develop a custom data entry program.

Sometimes the source document is a scanned image provided on-screen along with the data input fields; in other cases, we are asked to gather information from specific locations on the Internet; and in a few cases, the customer sends us PDF files.

In this model, the cost will not be affected by document handling, scanning, or programming.

The most critical factors affecting cost are the actual labor cost of data entry, start-up costs, and ongoing project management, as discussed above.

Concerning the actual data entry, in addition to the factors mentioned above, the design of the customer’s program can significantly affect operator productivity and hence cost. To facilitate faster, more accurate processing, we recommend as many field controls as possible. For example:

- Restricting fields to valid values and formats (e.g. date, text, numeric)
- Having look-up tables for repetitive values
- Ease of navigation
- System responsiveness over the Internet

Start-up costs will be similar, in that the project manager must learn the system and the business rules, develop training materials unless provided by the customer, train and monitor the staff assigned. In order to provide a reliable fee quote, it is essential that the project manager have sufficient hands-on experience.

In many cases, the customer performs a quality review and reports to us common errors; in other cases, the system allows us to perform the review and the customer is willing to pay for it.

As customers become increasingly concerned with security and confidentiality, this model is gaining significant traction. It is also less expensive, in terms of both outside vendor costs and internal staff requirements, and turnaround is instantaneous. We like it very much, because it allows us to do what we are best at – managing and performing data entry projects, large and small.