

## Automatic Document Identification and Classification Software



**Don Field**  
CEO

**CEOCFO:** *Mr. Field, would you tell us the concept behind Ephesoft and Smart Capture?*

**Mr. Field:** Our idea was to bring a technology that has been around for years for Fortune 500 companies to the larger mass market. We have taken common tools used for high-end enterprise companies and brought them to a useable level by the Fortune 50,000. We benefitted from using open-source components. Open source is a concept where technology has been developed and shared with other developers and integrators worldwide. We used those tools to build our product to compete very effectively with proprietary technology.

**CEOCFO:** *What does Smart Capture allow people to accomplish?*

**Mr. Field:** It does a few things. It identifies documents and classifies them. A common application we have is mortgage loan processing. If you have ever purchased a home you know that there are many different types of documents. With increasing regulation the last few years, there now can be over 200 different types of documents you have to complete when you buy a home. Without technology, a skilled worker must look at each document, identify what document it is and separate it from the other documents. They probably add a special barcode on it to make it easily identifiable to the computer. We eliminate all this manual effort by automatically identifying the document and classifying it. Now if there are 200 different documents in a batch of a thousand pages, our software can go through and identify the first six pages are the appraisal and the second two pages are proof of employment, etc. We can do that in seconds which saves labor and therefore costs as well as speeding up the process so a loan can be approved faster. We actually can produce more accurate classification with fewer errors than a human so it is a process for improving processes and getting higher productivity.

**CEOCFO:** *Are you identifying by document number or are you going immediately to content?*

**Mr. Field:** It does not use any document numbers. It is strictly content-based classification. In effect, we look at the words on a page and compare that to words we know are in each type of document. Doing that allows us to precisely identify the document based on the words, just like a person would. So it is strictly content based and it works. Besides identifying the type of document, it can then pull off any data on the document that you are looking for automatically. For example, in an appraisal, there might be three comparable properties that you are appraising against and we can pull up the value of those comparables plus the addresses, sold dates, sizes, etc. We can put that in a database for the mortgage loan officer to see. Instead of having to look at all these documents on paper, they can just look at the data and do not have to look at the documents. It saves a lot of labor for different people in the organization and it makes the process faster. We talked about mortgages but it works with insurance claims, manufacturing as well as accounts payable applications. It classifies and identifies documents and then it can pull up any data that the user wants.

**CEOCFO:** *Would you give us an example of a few documents that are similar in content and how you are able to identify one from the other?*

**Mr. Field:** There might be different revision levels on documents or there might be notes that say addendum or revision dates or things like that. We can read all of those and automatically know it is a different document. We also have logic that we can determine even if the documents are exactly the same but the dates are different. We can put logic in that says this one has a later date so it is the more current version. Once you can identify the words on the page, now you will have a lot of ability to apply some intelligence to it. It is the same thing a person would do. So if a person were to look at two similar documents, how would they know which is more relevant? They would look for signs on the documents whether the revision level was different or the date was different.

**CEOCFO: What were the biggest challenges in putting the system together?**

**Mr. Field:** Our biggest challenge for us as a company is visibility because we are a new company. We are about 5 years old and we all came from this industry so we know the product and we know the industry as well as how it fits. It is challenging however when you are talking to a sizeable company to have them take a chance on a smaller company and feel there is confidence that we can do what we say. The way we overcome that is by doing proof-of-concepts so where we prove the technology using their own documents. We show them how it will work for their documents and that normally does the trick.

**CEOCFO: You mentioned the technology has been available for larger companies. Are you able to reduce cost by the methods you are using?**

**Mr. Field:** Over the past decade, similar solutions using different tools were developed by other companies and by big companies like IBM and EMC. But their offerings are quite expensive so if you are a Fortune 500 bank, this would be something you would invest in and you could spend millions of dollars since they do have a strong ROI. The problem is these tools are owned by those companies and they are quite expensive. For a mid-tier bank to compete with Chase or Wells Fargo they may not have millions of dollars to spend on a similar system but they still have the same needs as the big banks. Our software using open-source tools brings the price way down because we did not have to develop the underlying tool set and we can pass along those savings. If we had to develop all of those tools from scratch, it would be quite expensive too. By using the open-source community, we are able to leverage that to provide software at a much lower price. But price isn't the main price to buy Ephesoft. It's the modern architecture, openness of the platform, scalability, and effectiveness of the solution. Also our delivery method is primarily web-based; we do not have a direct sales force since we use a network of certified systems integrators and channel partners. Therefore, our selling and marketing costs are much less as well as our development costs. Consequently, the prices can be just a fraction of what they would be through a major old-line company with a long legacy installed base.

**"Ephesoft is a technology that brings a hard dollar ROI to companies. Companies have either a document or an information problem. They have too much information and do not have it properly categorized for the data pulled from it. We make this process more effective and faster to enable cost savings and better business decisions." - Don Field**

**CEOCFO: Are companies skeptical that it can be done accurately or are they skeptical that you are able to do it at all?**

**Mr. Field:** It is a combination. Companies that are very cautious still have doubts that technology can solve their problems in general. I think that the pace of technology change the last ten years has been so great and the consumerization of technology has meant that for the first time users and companies have better tools and computers than even the company has. I think companies are finding they need to move faster. We have seen increasing openness to exploring new solutions to their business problems. I think in the last five years things have opened up to that. They are still very careful in their evaluation and they should be but I think they are more open to new technology these days. Everybody is under cost pressure so if you have an ability to have a hard dollar ROI in a short period like we do; it does get in the door to people even if they are skeptical.

We compete with companies like IBM so there is an old saying that you cannot get fired for picking IBM. It is a safe decision. Going with a smaller player who has been in business five years as opposed to fifty years, is a bit more challenging to get approved by higher levels of management in some companies. The way around that is to do thorough testing and proof-of-concept demonstrations with customer documents. When we compete with these other companies, we ask the customers to let us prove it to them. We will challenge anybody to do the same thing in the same amount of time. That has worked for us so we compete to do the same work, we generally win. The other thing is visibility; I would say an obstacle to faster growth is visibility so we did not focus much on marketing for the first few years. We just recently hired our first marketing person. We had a website and good contacts in the industry but we did not focus on marketing so now we are going to do that. We are going to more events and tradeshow and have a better website as well as spending money on search engine optimization (SEO) so people can find us. Once they find us, it is kind of a revelation; they all like our approach and the value in what we are doing. I think our biggest obstacle is being found and that is what we are working on now.

**CEOCFO: What is your global reach today?**

**Mr. Field:** We have customers in 27 different countries now. We have a strong presence in Europe and Africa and that is because they are more accepting of newer technology particularly open-source technologies. We are intending to be a worldwide company. The growth in Europe and Asia and Latin America is actually greater than the growth of America. At

this point I would say that about 60% of our sales are in the Americas and 40% in the rest of the world but it is quickly becoming 50/50.

**CEOCFO: *What is involved in implementation?***

**Mr. Field:** We decided not to become the direct provider of all the software. We use a network of systems integrators and resellers around the world. We currently have about 150 of these types of partners that are all certified by us by attending training and passing a test. So we know they are capable of properly showing and installing the product with confidence. If a customer today were to talk to us first, we would get them in-touch with the right type of partner and that partner would work directly with them and install the software for them.

**CEOCFO: *How much customization is needed?***

**Mr. Field:** For classifying documents it is pretty quick. I would say those types of projects are one or two weeks. When you want to do extraction of data which is pulling data off the form, that requires a little more work so we are talking about one or two months. The normal project includes both of those things. The normal project is about two months long and then the customer is ready to go into production.

**CEOCFO: *What has changed from your original concept for Ephesoft?***

**Mr. Field:** We thought the mid to larger companies would be more aggressive about finding lower cost solutions and we found them to be more cautious than the smaller companies. The smaller companies seem to be more eager to try something new. We have a whole mix of companies including large, middle and smaller.

**CEOCFO: *Why pay attention to Ephesoft?***

**Mr. Field:** Ephesoft is a technology that brings a hard dollar ROI to companies. Companies have either a document or an information problem. They have too much information and do not have it properly categorized for the data pulled from it. We make this process more effective and faster to enable cost savings and better business decisions. Our technology enables that to happen at a rapid pace compared to current processes. It is rare that somebody would not have less than a six-month ROI using our technology whereas it could take two to three years with our competitors and be slower to implement and at a higher cost. We think we are comparable to the major players in the industry in capabilities but at a fraction of the price.

Interview conducted by: Lynn Fosse, Senior Editor, CEOCFO Magazine

---



**Ephesoft, Inc.**

[www.ephesoft.com](http://www.ephesoft.com)

**Don Field**

**949-335-5335**

**[don.field@ephesoft.com](mailto:don.field@ephesoft.com)**