

Hello. It looks like you're using an ad blocker that may prevent our website from working properly. To receive the best experience possible, please make sure any blockers are switched off and refresh the page.

If you have any questions or need help you can email us

## TECHNOLOGY

# Scan, Organize... and Now Populate: The Workflow Revolution Continues

Just a few short years ago, the paperless office movement started a new class of programs designed toward creating a digital professional environment. While there may be some environmental advantages to the paperless office paradigm, the major proponents of the effort were far from tree huggers.

Sep. 01, 2008

*From the Sept. 2008 Issue*

Just a few short years ago, the paperless office movement started a new class of programs designed toward creating a digital professional environment. While there may be some environmental advantages to the paperless office paradigm, the major proponents of the effort were far from tree huggers. Instead, they were savvy professionals who understood the technological landscape and saw the many opportunities that such a digital revolution could provide.

These benefits included practical issues such as limiting the need for storage space, but they also allowed for more reliable document retention, instant access to client records and the ability to keep all client-related documents together by creating digital folders. As the technologies evolved, additional features became available, enabling professionals to create fully electronic digital workpapers that enhance collaboration between professionals or between the client and the firm, as well as offering dynamic links (drill-down functionality) from returns, financials and other work product to the supporting documents,

spreadsheets,  
programs and other related materials.

Revolutions are not pain-free, of course. This substantial change caught many firms off-guard as they implemented solutions that required them to redesign their internal workflows. And when looking at the greater picture, the paperless movement was only a part of the larger issue of revamping firm processes to take full advantage of the productivity and efficiency-boosting capabilities of these new programs.

The most recent innovation in the workflow revolution is automated population of 1040 tax returns with data from source documents. Ever since the concept was first touted a few years ago, it has been generally embraced as a way to quickly eliminate much of the labor involved in data entry. But with optical character recognition (OCR) not quite reliable enough for truly hands-off automatic transfer of data into forms, technology vendors in the tax and document management spaces had to develop workflow methods that would optimize the automation aspects of the concept, while instituting efficient review processes.

Through several years of development, several products are now coming to the market. Much more than just add-ons to a tax system, these automated population utilities bridge between a document management application and a professional's tax program. In most cases, the systems are offered as a part of a vendor's document management suite, essentially making these systems into a new breed of programs that "scan, organize and populate."

How the Process Works Each of the scan, organize and populate systems on the market works in different ways and has varying workflow processes, but at their base, they are designed for a firm that has a front-end scanning process in place.

In other words, the first thing to do when a client delivers a stack of documents is to put them in a scanner. Since the quality of the scanned document is greatly affected by the quality of the original, it's best to scan the originals, not photo copies. After this, the systems manage the scanned documents in different ways, generally providing an electronic work folder or other filing system.

The systems then look at each form to determine what they are (W-2, 1099, K-1, brokerage statement, etc.). This is the hard part, and it's the area in which most of the vendors concentrated since there are literally thousands of slight variations to even the most common forms — think about all of the different bank-produced 1099-INTs out there. The better a system is at recognizing these forms, the better job it does in the next step, which is pulling data from each of the fields and classifying it appropriately (as income, interest or other categories).

The programs currently on the market generally approach this form recognition function in one of two ways — either by having a built-in library of form types that the system uses to identify scanned forms, or by securely transmitting the scanned image format (not the data) to the vendor's large array of servers, which learn from the experiences of all users of the system. This latter method is much more robust, providing the most comprehensive recognition capabilities.

Once the form has been recognized and the data extracted, the next step is critical: verification of the data that has been scanned, after which can be transferred into a client return. This front-end check of the data is necessary to ensure accuracy. Once the data has been input into the return, it is editable just as any other entry and, depending upon the tax and auto-populate program, often offers drilldown access to the scanned document.

Of the half dozen systems on the market, most are currently being offered by developers of tax software, with the populate features only available for users of their professional tax preparation systems. However, two workflow automation vendors, SurePrep and Copanion, offer independent systems that integrate with several tax packages.

---

**CCH,**  
**a Wolters Kluwer business – ProSystem fx Scan with AutoFlow Technology**

[www.tax.cchgroup.com](http://www.tax.cchgroup.com)

The ProSystem *fx* Scan system, which was originally called BOCDIP and which was a recipient of a **2006 Tax & Accounting Technology Innovation Award** ([www.CPATechAdvisor.com/go/1196](http://www.CPATechAdvisor.com/go/1196)),

was one of the first breed of document recognition scanning systems. The program essentially works by allowing users to scan in all of a 1040 client's tax-related forms, statements and other documents, with the system using optical character recognition technology (OCR) to identify what the individual documents are and then produce a fully bookmarked and logically organized PDF file that includes all of the documents.

This results in a fully digitized set of workpapers that can be used in a dual-screen setup to review a client return. Bookmarks can be sorted to meet firm workflow needs, and the program includes strong search functionality for retrieving specific forms within client records. Since the program is part of the ProSystem *fx* Suite, it also offers integration with CCH's tax, write-up and document management systems, as well as with the vendor's client collaboration portals. Through the optional PDFlyer module, users are also given enhanced PDF review capabilities, such as the ability to move pages and bookmarks within the document, utilize tickmarks and connectors, or resize and rotate document views.

For the 2008 tax year, CCH has added the optional AutoFlow Technology system for ProSystem *fx* Scan. This new feature provides automatic population of client data into 1040 returns prepared using ProSystem *fx* Tax. The process is straightforward, with the program extracting data from previously scanned client documents, and then providing two validation features to ensure data accuracy. Users, generally administrative staff or junior professionals charged with scanning and inputting initial data, can look over OCR results and also ensure that items are classified properly. When the tax system is opened, the user is alerted that information is available for that client, and the user can then import it directly into the client return much like organizer data.

Both Scan and its AutoFlow system are locally installed on the firm's computers, which keeps all data in-house and generally provides for more intuitive ease-of-use, but can result in slower processing times when compared to systems that securely route form data to large processing centers with larger server arrays performing the work. CCH does offer some of the benefits of the server-side technology, however.

“During tax season, when time is at a premium, AutoFlow Technology will shift 1040 data entry functions away from a firm's professional staff,”

said Bob Dias, CCH Vice President of Product and Segment Management. “Any solution that saves significant professional time is obviously of huge benefit.”

AutoFlow Technology is imbedded in Scan v4.0 and available to all ProSystem *fx* Scan users without an additional license fee. A per-return transaction fee is applicable for importing the extracted information into ProSystem *fx* Tax. The list price per return is \$15, irrespective of the number of imports or the number of forms imported. The \$15 authorization list price has been reduced to \$10 for an initial promotional period.

**Forms Supported for Automatic Population:** W-2, W-2G, 1098, 1098 T, 1099 DIV, 1099 INT, 1099 MISC, 1099 Q, 1099 R, 2439, SSA 1099 On the development roadmap: Forms K1-1065, K1-1041 and K1-1120S, Combined Statements.

---

## CCH

### Small Firm Services – ATX Scan&Fill & TaxWise Scan&Fill

[www.ATXinc.com](http://www.ATXinc.com); [www.TaxWise.com](http://www.TaxWise.com)

Since the acquisition of ATX and TaxWise by CCH a couple of years ago, the two tax preparation systems, both geared toward small and mid-sized firms, have seen increased development, thanks to the considerable resources that CCH has to offer. One of the more notable developments over this time is the Scan&Fill module, offered under each of the tax systems' names. The module, which includes a basic Document Manager system, provides direct scanning using most scanners (sheet-fed is preferable), with optical character recognition (OCR) extracting client data from W-2s and most versions of 1099.

Operating separately from the tax system (either ATX or TaxWise), Scan&Fill automatically creates a temporary client folder that stores documents that have been scanned, with items sorted by document type. Each of these documents is saved as an individual image file (\*.jpeg), while the system's Review function provides a panel that displays the data that has been extracted from an individual form and allows editing as necessary.

Scan&Fill's AutoFile feature automatically files recognized documents in the client's folder or the program will automatically create a new client folder and file the document if the client is not yet in the system.

After the user has verified the accuracy of the data in the scanned form, the next step is to move the temporary folder into the Active Client Directory, which is part of the document management features of the program and provides separate folders for source documents and the client tax return. Individual folders for spouses filing jointly can be linked so that when the user chooses to populate the return, data from both are included. After these review and folder management tasks, the process of transferring client data is performed by exporting the data into a file that is then imported by the ATX or TaxWise preparation system.

While preparing a tax return, users can access scanned source documents and, when the return is completed it can be stored as a PDF with the client's documents in the document management module. The Document Manager also enables storing of other digital files and documents related to a client engagement, such as signature pages, receipts and notes. It also provides basic search functionality.

Although its form recognition and workpaper management capabilities are less comprehensive than many of the other auto-population systems designed for larger practices, Scan&Fill's feature set is well suited to smaller and mid-sized practices using TaxWise and ATX and who are looking for a paperless solution that also provides time savings by automating much of the data entry associated with 1040 tax preparation services. The system costs \$715 for a single-user license, \$1,035 for up to five users.

**Forms Supported for Automatic Population of a Tax Return and AutoFile to the Client Folder: W-2, 1099 (most versions)**

**Additional Forms Recognized for AutoFile to the Client Folder:**

K-1, 8879

---

**Coppanion**

– GruntWorx

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

Coppanion is a new entry into the space, having debuted GruntWorx last fall. The fully web-based workflow management system provides scanning, identification



and bookmarking of all client documents into a set of digital workpapers that enables users to quickly jump to any document when reviewing a return or performing other tasks. The system also allows the addition of tickmarks and notes from multiple users, allowing firms to keep much of their traditional processes intact. In its first year (TY2007), professional users at firms ranging from the Big 4 to small practices processed more than 350,000 client documents using the system.

In time for next tax season, Copanion is adding automatic tax return population features to the Pro version of GruntWorx, making it one of the few products on the market that is not from a tax software vendor and that integrates with multiple systems. For the 2008 tax season, GruntWorx Pro will automatically export data from scanned tax documents into GoSystem Tax RS. Support for additional tax preparation software packages will be determined and announced later in the 2008 tax year.

Among the more notable features of GruntWorx are several patent-pending technological aspects, including biometric recognition of forms. This feature in particular, which is similar to the technique used by law enforcement for analyzing fingerprints, gives the system the ability to recognize the most minute differences between various forms. Couple this with the intelligence of the program, which can actually learn as it goes and share this experience across all users, and the result is that GruntWorx has the ability to recognize virtually any type of form or consolidated statements, regardless of the issuer. The system also has the ability to identify tax organizers and handwritten notes.

The resulting product is an organized, bookmarked PDF file and coversheet, with a summary of client data that can then be uploaded into one of the supported tax preparation systems. GruntWorx also has an Image Enhancement feature, which improves the quality of scanned input. All of this processing power gives the system extraordinary accuracy, but requires significant infrastructure, which Copanion maintains at its secure facilities.

The actual user process is very simple, and the initial functions take only three steps:

1. Select a client,
2. Scan client documents or add documents from previously scanned PDF files, and
3. Upload the data through the secure GruntWorx system.

GruntWorx then processes the files and identifies, classifies, sorts and compiles the bookmarked PDF workpapers. When processing is complete (usually in one or two hours), the program notifies the user, who then securely downloads the PDF. Tax professionals can use the bookmarked PDF to quickly find and enter tax data in any tax software. GruntWorx Pro provides the additional capability to extract data from the scanned tax documents and automatically populate the data in tax preparation software. The extracted data can be verified for accuracy against the scanned source documents in the bookmarked PDF.

All data is secure, remains in the United States and is never accessible to outside eyes, making the system fully compliant with IRC 7216. Pricing for the scan & organize function is very affordable at \$1 to \$2 per client based on volume, regardless of the actual number of documents and pages scanned. Pricing for GruntWorx Pro for GoSystem Tax RS will have a list price of \$50 per return. “Automatic population of client data can save significant administrative time, but there are many challenges that must be overcome to develop a reliable system,” said Ed Jennings, Copanion’s vice president of sales and marketing.

“There can be thousands of variations of even the most basic forms, depending on the issuer, so we focused on ensuring the accuracy of classification across a broad range of forms. We also wanted to make sure the system was intuitive and easy for users to understand.”

**Forms Supported for Automatic Population:** GruntWorx can classify and extract data from virtually all forms (1099, 1098, W-2, K-1, 2439, SSA, etc.) and combined brokerage statements.

---

**SurePrep**

— 1040SCAN

[www.1040SCAN.com](http://www.1040SCAN.com)



1040SCAN is the most mature digital document organization and auto population system on the market, having been used the past three tax seasons by thousands of professionals. For TY2007, the secure, web-based system and its population features were used to process more than 125,000 returns, and the company expects that number to double next tax season. SurePrep also offers a Lite version of 1040SCAN that offers only scan and organize (no population) features, which is best-suited to users of non-supported tax programs. The vendor also offers domestic outsourced tax preparation services.

The key benefits to 1040SCAN are its population capabilities, which integrate with GoSystem Tax RS, ProSystem fx Tax, and Lacerte to automatically extract data from scanned documents for preparing client 1040 returns. As a part of this process, the system automatically identifies all of the scanned documents and organizes them, then provides the user with a bookmarked PDF file that includes all of the items in an order that follows the flow of the return. This set of electronic workpapers can greatly streamline the review process, with documents easy to find and quickly retrievable. 1040SCAN can identify and extract data from thousands of variations of all 1040-related forms and consolidated statements from more than 70 brokerage entities.

In great part due to its maturity on the market and the fact that the company has more than 200 professional U.S. tax preparers on its staff who help in program development, 1040SCAN provides the most advanced data verification and analysis tools. In fact, SurePrep built a separate diagnostics program focused specifically on classifying documents and identifying potential duplicates or errors. The result is much greater accuracy than simply relying on OCR and can alert users to issues such as data previously entered on an organizer that has also been scanned into the system. The program also offers a series of diagnostic reports.

Operation of 1040SCAN is simple and geared toward non-professional administrative staff who first scan the client documents and upload them to SurePrep's secure servers, where they are identified, organized and data is extracted for use by the population feature. The next step is completion of the 1040SCAN Review Wizard, a verification process that guides the user (ideally still an admin or intern) through a checklist of potential errors or duplications. The bookmarked PDF is available after this process, and client data can be exported into one of the supported tax systems with just a few mouse clicks. The professional

assigned as the preparer is then notified that this client return is in progress and awaiting action.

According to Bret Wier, SurePrep's VP of Sales and Marketing, optimizing a firm's workflow is one of the biggest factors in achieving the greatest results from such a system. Additionally, the practice should utilize up-front scanning and data verification and use a high-quality scanner.

"The quality of the scanned images is critical because it affects the ability of the OCR to recognize forms and data contained on them," he said. "Each year, we stress to our users that having a good scanner is essential, and we often recommend the Fujitsu 6130 with VRS from Kofax. That scanner model costs about \$1,100 (from Dell.com), and Wier notes that it is a good fit for firms preparing up to 2,000 client 1040 returns.

1040SCAN also includes numerous workflow management tools, enabling preparers and managers to quickly see the progress of returns assigned to them, and alerting them to when action is required on their part. These tools are especially valuable since they help firms implement more efficient workflow processes that can help save as much as 25 percent of senior reviewer time. 1040SCAN's list price is \$30 per return with a 20 percent early season discount if purchased prior to the end of September 2008.

**Forms Supported for Automatic Population: W-2, W-2G, 1099**

**B, 1099 C, 1099 G, 1099 DIV, 1099 INT, 1099 LTC, 1099 MISC, 1099 OID, 1099 PATR, 1099 Q, 1099 R, 1099 SA, 1098, 1098 E, 1098 T, SSA-1099, 5498, MA 1099-HC, K-1 (1065, 1120s, 1041)**

- K-1 (Resident & Non Resident: CA, IL, MA, NJ, NY, OH, PA, VT, WI)
- Brokerage statements from more than 70 firms
- Grantor letters from major financial institutions
- Also adding 1041 returns for next year

GoFileRoom ES is a secure, web-based document management system offered as part of the Thomson Reuters Enterprise Suite family of high-end, professional accounting and tax compliance programs or as part of the CS Professional Suite, which is geared toward smaller and mid-sized practices. Both suites offer a variety of applications for professional accounting firms, ranging from tax and write-up to practice management and planning. While the features and capabilities of the ES and CS flavors are exactly the same, the two offerings differ in integration and pricing.

GoFileRoom ES offers extensive document management and workflow features, organizing documents in an index-based structure that enables thorough search and query functions, as well as offering full linking of source documents to and from client returns. Through the program's FirmFlow module, users can manage all aspects of client engagements, with management of all scanned documents, files, notes, checklists and firm-specific routing slips or other items. With the addition of the TaxSort module about two years ago, GoFileRoom ES also offers advanced organization functions, with the web-based system processing scanned 1040 source documents, identifying what each document is, and producing a bookmarked PDF file that organizes the documents into the correct tax preparation order, along with a summary cover sheet of extracted client data.

Since the core features of this module rely upon recognition of forms for which there can be numerous variations, the program is specifically designed to learn from previous use. That is, the first time a user at a firm scans in a particular form that has never been identified by the program, which is housed on Thomson Reuters secure servers, it might take a little longer to process. But the next time someone at that firm or any other firm using GoFileRoom ES submits that particular form, the recognition process is almost instant. No client data is retained in Thomson Reuters servers, only basic information about the formatting of documents, which helps it build its knowledgebase.

GoFileRoom ES can recognize virtually all common forms and even infrequently used ones, as well as combined brokerage statements of any length from almost all sources. The system integrates directly with the GoSystem Tax ES compliance

package, but it does not currently offer an automatic client return population feature. Frank Swierz, a Senior Director of Software Development for Thomson Reuters, noted that they are exploring development of such a system, as well as considering partnering with third-party developers for integration with their auto populate capabilities.

There is no separate module fee; per-return pricing for TaxSort starts at \$1 per tax return, and the per-return price decreases as volume increases.

**Forms Supported for Automatic Population:** None at this time, although Thomson Reuters is considering development of an automatic population feature in the future (see above). On the CS side, the FileCabinet CS solution does offer a population module with integration into UltraTax CS.

---

**Thomson  
Reuters — FileCabinet CS**  
[cs.ThomsonReuters.com](https://cs.thomsonreuters.com)

FileCabinet CS is part of the Thomson Reuters CS Professional Suite. Initially designed as a paperless document storage system that houses scanned client documents

as images along with digital document files in a basic folder-based format, FileCabinet CS has evolved over the past few years to include various organization and management tools, including the ability to add annotations to documents, track histories, integration of e-mail and fax functions, as well as reporting and analysis features. Thomson Reuters also offers GoFileRoom ES, a more robust document management system that provides digital workpaper output.

The most recent additions to FileCabinet CS have focused on workpaper organization

and the addition of optical character recognition (OCR) capabilities that enable automated population of client 1040 returns processed using the UltraTax CS professional preparation system. This year, Thomson Reuters debuted the optional Source Document module, which offers form recognition, extraction of form data and the ability to integrate with the tax package. This new module is specifically geared toward use of the automatic population feature, since it does not result in compiled, organized and bookmarked PDF or other advanced digital workpaper output, but instead stores the data in a client folder as individual items.

While FileCabinet CS is generally locally installed (although it can be remotely hosted), the Source Document module is a web-based feature through which users scan documents and securely transmit them to the Thomson Reuters Data Center for OCR processing, data extraction and page naming. When returned to the user, the extracted tax data can be reviewed (best done with a dual monitor), edited using the UltraTax CS Source Data Entry utility if necessary, and then imported into the tax system.

For users familiar with any of the programs in the CS Professional Suite, navigation will be familiar, and the entire process should be fairly intuitive, with only a few steps to follow. First, documents are scanned and grouped, then placed into a folder; they are then sent to Thomson Reuters for processing. When processing is complete, the user receives individually named PDFs for each document and is notified via the Message Center in UltraTax CS that client data is available for entry into the system. After verifying the data, the user can select the populate function, which routes data to appropriate entry fields.

Since the program is web-based instead of locally hosted, it offers the advantage of being able to learn form formats and other information (not client data), which enables it to operate increasingly more efficiently and quickly as tax season progresses. Additionally, as a system designed for firms using an up-front scanning process, it frees professional staff from the administrative tasks associated with document scanning and management.

There is no separate fee for the Source Document Module; however, it will have a per-return pricing fee, which was yet to be determined at the time this article was written. FileCabinet CS starts at \$1,500.

### **Forms Supported for Automatic Population:** FileCabinet CS

with the Source Document module can be used to classify and extract data from numerous forms (W-2, 1099-INT, 1099-DIV, 1098, 1099R, 1099MISC, 1099G, 1099B, W-2G, 1099Q, 1098E, 1098T, 2439 and SSA).

---

It is important to remember that automated population of client data is only one step in the new workflow processes evolving to make professional practices more efficient and productive. Just as technology has affected almost every aspect of the profession, it also requires that firms be flexible and open to new workflow methods.

New technologies such as scan, organize and populate systems may have the potential to greatly optimize productivity, but adaptation of work processes is necessary to get the most out of them. It's also wise to have the right hardware: Dual-screen monitors are pretty much essential, and firms should invest in a high-volume scanner or a workgroup model (see [www.cpatechadvisor.com/go/2116](http://www.cpatechadvisor.com/go/2116) for more information about scanners).

Technology • Article

CPAPA is registered with the National Association of State Boards of Accountancy (NASBA) as a sponsor of continuing professional education on the National Registry of CPE Sponsors.

© 2023 Firmworks, LLC. All rights reserved