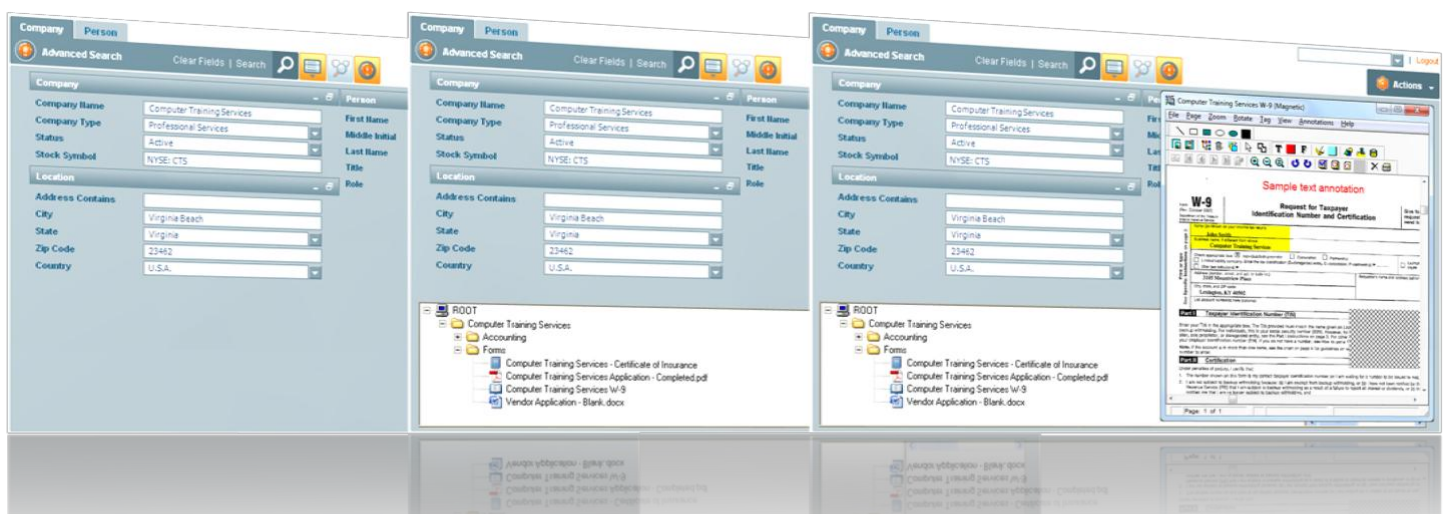


How Embedded, Private-Labeled Document Management Adds Value to Your Existing Software

(A White Paper for ISVs and Other Solutions Providers)



Transform your "data-centric" software application

into a "content-enabled" software solution

in less than a day using the Datahaven SDK for ECM.



TORRENTIAL DATA
Bridging the Content and Data Divide

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
DATA-CENTRIC VS. CONTENT-ENABLED: AN EXAMPLE	2
HOW CONTENT-ENABLED SOLUTIONS BENEFIT YOUR CLIENTS.....	3
<i>GENERATING SIGNIFICANT HARD DOLLAR COST SAVINGS</i>	3
<i>PRODUCING DRAMATIC GAINS IN PRODUCTIVITY</i>	3
<i>ADOPTING BEST PRACTICES & IMPROVING LEGAL COMPLIANCE</i>	4
<i>LEVERAGING EXISTING IT INVESTMENTS</i>	4
HOW CONTENT-ENABLED SOLUTIONS BENEFIT SOLUTIONS PROVIDERS.....	4
<i>MEETING A CRITICAL NEED OF YOUR CUSTOMER BASE</i>	5
<i>DELIVERING SOPHISTICATED FEATURES IN LESS TIME USING FEWER DEVELOPERS</i>	5
<i>ENHANCING LICENSE & SERVICES REVENUE</i>	6
CORE ELEMENTS OF A DATAHAVEN-POWERED SOLUTION	6
<i>EMBEDDING A SCALABLE, ENTERPRISE-CLASS CONTENT REPOSITORY</i>	7
<i>EMBEDDING A CONTEXT-SENSITIVE VIRTUAL FILE CABINET</i>	8
<i>LEVERAGING OCR & OTHER AUTOMATED DATA EXTRACTION</i>	10
<i>EMBEDDING CONTENT-CENTRIC WORKFLOW COMPONENTS</i>	10
<i>ACCESSING CONTENT (NOT SEARCHING FOR IT)</i>	11
<i>SECURING ACCESS TO CONTENT & AUDITING USER ACTIVITIES</i>	12
CONCLUSION	13

EXECUTIVE SUMMARY

Does your organization develop custom, enterprise-class software applications or customize popular third party enterprise applications such as Dynamics, PeopleSoft, SAP, and Salesforce?

Most software manages 20% or less of an organization's information

Do you realize that the “structured data” managed by these applications—things such as names, dates, and transaction amounts—represent 20% or less of an organization’s total information. In other words, 80% or more of an organization’s total information is found in the form of “unstructured content” such as paper, faxes, emails, and other electronic files. And upwards of 90% of this unstructured content is left unmanaged.

A recent survey by IDC reflects the mounting frustration of information workers with this reality. A significant majority of these workers now expect an integration of enterprise content management (ECM) technologies with data-centric enterprise applications such as ERP and CRM systems “so they can control their growing volumes of unstructured information in the context of their enterprise business processes.”

Software users want to see vendors work together to create pre-integrated solutions that manage structured and unstructured information

Despite the fact that nearly 60% of these respondents “would like to see vendors work together to provide pre-integrated solutions,” IDC concludes that few vendors of enterprise applications have delivered. All of this presents ISVs and other solutions providers with the opportunity to deliver significant value to their client base and provides a source of substantial additional revenue.

From document imaging to email and fax integration, OCR data capture, workflow, and content publishing and distribution, Datahaven provides the means to enable a data-centric application to natively capture, process, store, and manage the entire spectrum of content and related metadata from within its already-familiar user interface.

This provides information workers access—from directly within the applications they rely on every day to do their job—to precisely the right content at the precisely the right time based on their identity, security permissions, and the context of their task.

WHAT YOU NEED TO KNOW ABOUT DATA AND CONTENT

- Data-centric software focuses on the data of a record without easy access to the 80% of information in documents, faxes, and email.
- Finding information in documents, faxes, and email is difficult, time-consuming, and expensive and adds compliance risks.
- Simply improving the “findability” of this information would save a large enterprise approximately \$10 million per year.
- Content-enabled software provides instant access—natively from the already-familiar UI of the data-centric application—to the complete record, including related paper, faxes, and email.
- 60% of information workers want vendors to work together to provide packaged integrations of data-centric applications and document and content management.
- File attachments, low-level SDKs, and stand-alone systems fail because they can be too simplistic, too expensive, too complex, and/or too bloated.

DATAHAVEN OFFERS A UNIQUE SOLUTION

- Embed components for document and content management directly into an existing UI of COTS, custom-written, and popular third-party software applications that information workers use every day.
- Leverage a scalable, enterprise-class repository.
- Virtual file cabinet provides programmatically-filtered, context-sensitive views of all forms of content.
- Employ barcode recognition, OCR and other technologies to minimize data entry and human error.
- Automate processes using document-centric workflow from within the existing Dynamics UI.
- Enable users to access all relevant content from within the existing Dynamics UI instead of searching for it in a separate application.
- Enforce granular security settings and track all actions in audit trails.

DATA-CENTRIC VS. CONTENT-ENABLED: AN EXAMPLE

The most common examples of data-centric software applications are CRM and ERP systems. Data-centric software, however, can be just about any application that fails to provide information workers with access to all relevant documents and other content from directly within its existing user interface. Figure 1 provides a screenshot of a simplified example of a data-centric application.

Documents and other content related to the transactions and other records managed by data-centric applications typically will be stored using on-site and off-site paper storage, the organization's public folders, network file share and fax server, an individual's personal email archives and, in some cases, as files randomly attached to individual records somewhere within the existing data-centric application. No matter where these documents and other content might be stored, it's difficult, time-consuming, and expensive for information workers to find the specific item of content they need when they need it.

By content-enabling a data-centric application, solutions providers can offer support for functionality such as document imaging, email and fax integration, content-centric workflow, publishing, distribution, and the ability to natively manage any form of content and related metadata—all from directly within the existing user interface of the data-centric application. As a result, an information worker using a content-enabled software solution has instant, security-controlled access to, and the ability to work with, precisely the right content at precisely the right time based on the worker's identity, role, and the task at hand—all from directly within the familiar user interface of the application they use every day.

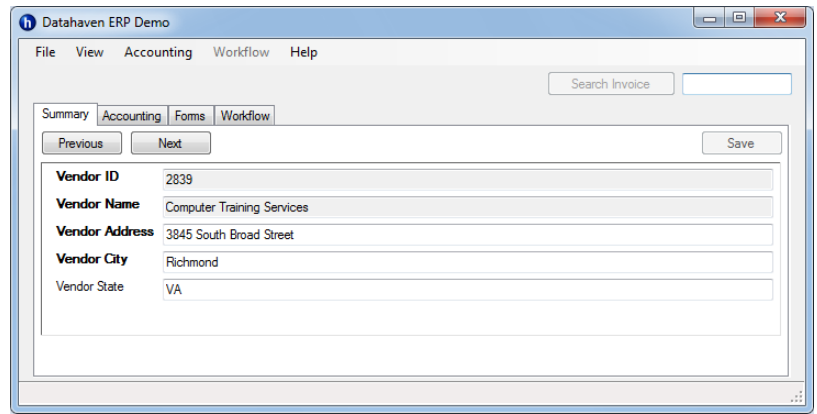


Figure 1: A simplified example of a data-centric application that fails to provide its users with easy access to documents and other content related to its records

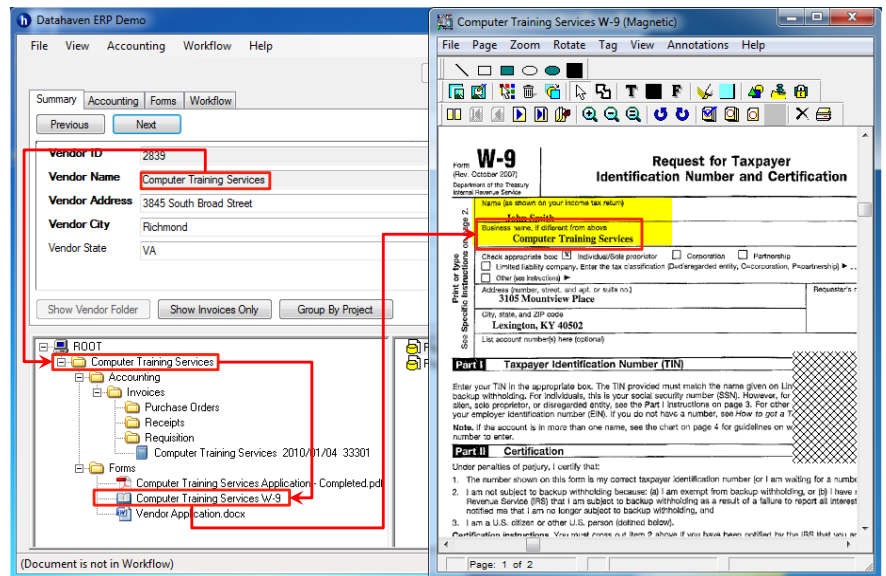


Figure 2: A content-enabled version of the same application in Fig. 1 that uses Datahaven to display a programmatically-controlled folder structure containing all relevant content related to the current record and a viewer that shows a highlight annotation and security redaction

HOW CONTENT-ENABLED SOLUTIONS BENEFIT YOUR CLIENTS

As a solutions provider, you know that you have to deliver value to your clients. The value and intangible benefits of the content-enabled solution you create arises in a number of important areas for the end user.

Generating Significant Hard Dollar Cost Savings

Generally speaking, ECM—enterprise content management—technologies have demonstrated the ability to generate substantial savings on expenses for administrative labor, on-site and off-site storage, and priority shipping. Embedding ECM technologies to create content-enabled solutions (rather than requiring users to use independent, stand-alone ECM systems) has the added benefit of saving substantial training costs associated with complex, stand-alone ECM systems because users of a content-enabled solution can intuitively find, and otherwise work with, any related content from within their data-centric application’s existing user interface.

Producing Dramatic Gains in Productivity

Implementing advanced ECM technologies enables an organization to adopt repeatable, reusable processes, minimize human error, and maximize information worker productivity. In the absence of ECM technologies, these information workers must continue to rely on primitive tools such as personal email archives, public folders, manual file attachments, and physical file folders to help them manage the content relevant to their daily tasks.

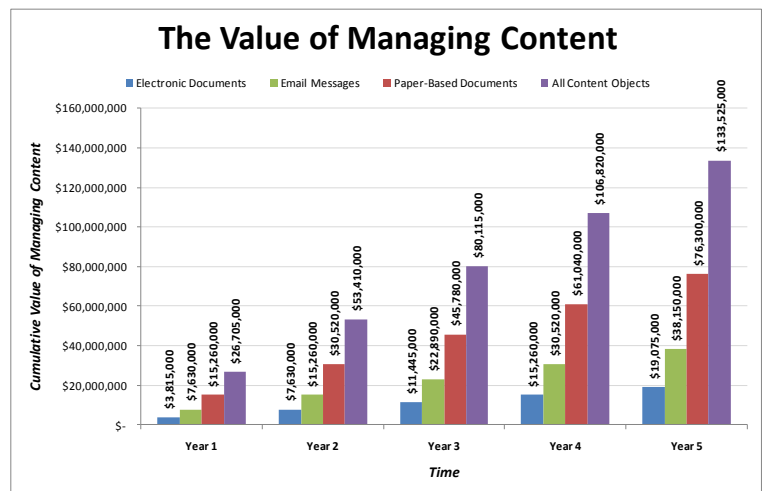


Figure 3: The cumulative value of managing content (or cost of failing to manage it) in an organization of 1,000 information workers

According to IDC, the average information worker spends in excess of 15% of the work week—that’s more than six hours every week—unsuccessfully searching for the documents, faxes, email, and other content needed to do their job. IDC concludes that simply improving the “findability” of the information would save an enterprise with 1,000 knowledge workers approximately \$10 million per year in worker productivity.

AIIM, the Association for Image and Information Management, estimates that organizations without ECM technologies will, on average, misfile 7% of the content they handle and incur a cost of \$128 to find each misfiled item. AIIM also estimates that these organizations, on average, lose 3% of their content and incur a cost of \$210 to re-create each lost item.

Using AIIM’s cost estimates and assuming each of IDC’s hypothetical 1,000 information workers handle an

average of 250 electronic documents, 500 emails, and 1,000 paper-based documents each year, we can see how IDC's hypothetical organization could reach \$10 million per year in savings. Of the 1.75 million items handled each year by this organization, the organization will misfile 122,500 of them and lose 52,500 of them (see Figure 6, p. 7) at an annual cost of more than \$26 million (see Figure 3, p. 3).

Adopting Best Practices & Improving Legal Compliance

Embedding advanced ECM technologies such as audit trails, document-centric workflow, and automated naming and filing methodologies directly into the user interfaces of existing applications will lead to improved legal compliance. Content-enabled solutions provide your customers with the ability to use these applications to enforce repeatable, reusable processes around document-centric processes within their organization and extend disaster recovery and business continuity plans and other best practices with structured data to the larger, more complex world of unstructured content.

Leveraging Existing IT Investments

A content-enabled solution functions as a seamless private-labeled version of an existing enterprise software application. This leverages existing IT investments and minimizes duplicative vendor relationships. Vendors of stand-alone ECM systems architect their applications and build their business models around the premise that managing documents and other content is best accomplished independently from the transactions or other daily tasks to which this content relates. They only integrate "after-the-fact. Therefore, these vendors require their (often bloated and always expensive and complex) application to be installed and used side-by-side with other enterprise applications.

HOW CONTENT-ENABLED SOLUTIONS BENEFIT SOLUTIONS PROVIDERS

As a solutions provider, transforming data-centric applications into content-enabled solutions offers you some unique opportunities, including the ability to broaden your client relationships and deliver substantial incremental value to your client base. Each of these opportunities can help you generate significant revenue growth.

Meeting a Critical Need of Your Customer Base

According to Gartner, end-users in the market for ECM technologies “express more of a tactical interest in buying something that is closer to commercial off-the-shelf (COTS) software than a framework or template that requires lots of customization.” Research from IDC reaches a similar conclusion: “58% of respondents from large organizations agreed/strongly agreed with the statement ‘we need better integration between our ECM and enterprise applications, and would like to see vendors work together to provide pre-integrated solutions.’”

Q. Please indicate your level of agreement with the following statements:

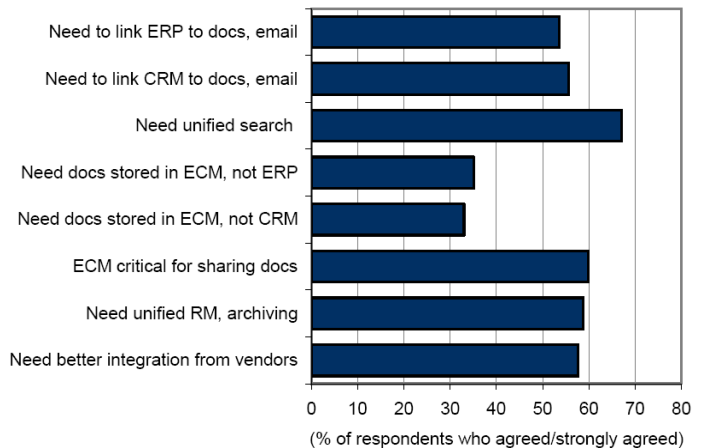


Figure 4: Customer perceptions of the need to integrate enterprise content management and data-centric enterprise applications (Source: IDC)

Perhaps this is why the major ECM suite vendors offer their own industry-focused content-enabled solutions that they have built on top of their ECM technologies (e.g., EMC Documentum in pharmaceuticals, FileNet in insurance, and Open Text in energy just to name a few). As Gartner points out, however, simply because these ECM vendors pursue a vertical market “does not necessarily correlate to deep functionality and domain expertise.”

Delivering Sophisticated Features in Less Time Using Fewer Developers

In the past, meeting this need of your client would require the client to adopt a stand-alone ECM system (and pay for a custom integration) or have a large enough project so you could be paid to become an ECM expert and build a custom solution essentially from scratch. Fortunately, solutions providers and their customers have a better alternative.

Embedded ECM components save solutions providers a substantial amount of time that would be spent

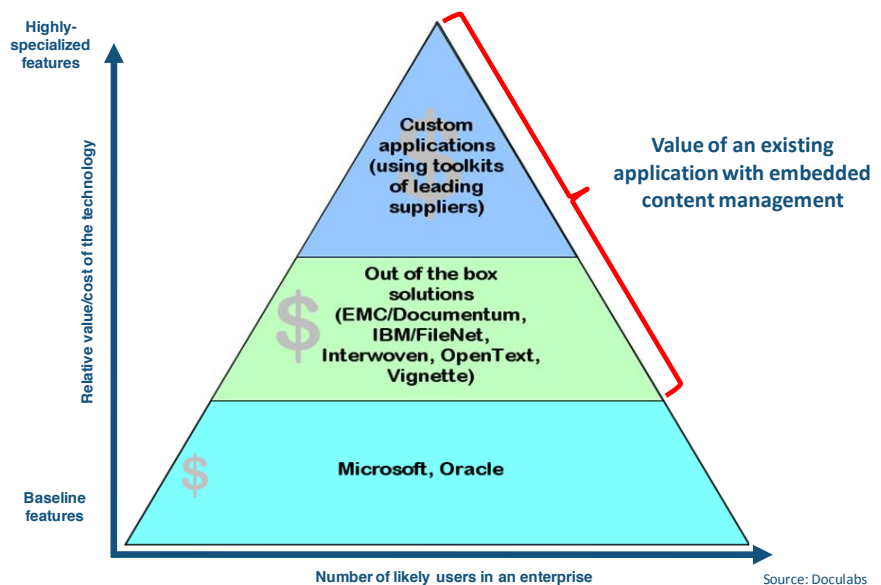


Figure 5: The relative value of content-enabled software vs. stand-alone ECM systems

- researching the capabilities of, and negotiating licensing terms with, multiple low-level SDK vendors for scanning, OCR, workflow, annotations, etc.
- recognizing and resolving the complexities involved in developing a document and content management solution
- coding and testing the interoperability among each low-level SDK and between each low-level SDK and your data-centric application
- ensuring that the functionality is presented to the end-user with a consistent look-and-feel throughout the data-centric application

The list goes on. The amount of time invested can quickly add up to multiple man-years of effort, delay milestones, and impact profitability. In some cases, the costs of including the functionality in a project may preclude it from ever proceeding past the planning phase.

With private-labeled ECM components, you can use a high-level toolkit to embed complete, advanced features in minutes—not months or years—without distracting scarce resources. This not only reduces a solution provider’s upfront development costs and ongoing expenditures for your internal support structure for a complex technology, it accelerates your time to market and allows your developers to focus their efforts on adding value to the core business application—not designing, coding, and testing its underlying infrastructure.

Enhancing License & Services Revenue

As depicted in the Fig. 5, the highly-specialized nature of a content-enabled software solution represents one of the most valuable segments in the technology industry. Solutions providers can generate recurring maintenance and support revenue in addition to substantial consulting services revenue related to analyzing and implementing each clients’ document-centric workflow processes. In addition, solutions providers have the opportunity to enjoy higher software margins in a market where, according to BuyerZone.com, entry-level, stand-alone ECM systems typically sell for \$10,000 (or more) for a handful of users, mid-level systems for about 100 users run from \$40,000 to \$100,000, and enterprise-level solutions for hundreds of users exceed \$500,000.

CORE ELEMENTS OF A DATAHAVEN-POWERED SOLUTION

Information workers continue to rely on powerful, data-centric technologies—including custom and COTS applications for ERP, CRM, and SCM—to perform their daily tasks. The problem is that each of these applications and each of the tasks performed by their users also produces (or requires information workers to process or reference) dozens of documents, faxes, and emails. This “content problem” is made worse by the ever-growing velocity, volume, and variety of the transactions and other “structured data” managed by these applications. Unfortunately, the typical approaches used to address this “content problem”—file attachments,

low-level SDKs, and stand-alone ECM systems—fall short because they can be too simplistic, too expensive, too complex, and/or too bloated.

Datahaven solves the “content problem” by enabling the applications that information workers use every day to manage—natively from within their already familiar user interface—the documents, faxes, emails, and other content related to the transactions and other records at the core of these data-centric applications. By creating “content-enabled” versions of these applications, information workers have instant access to precisely the right content at precisely the right time based on their identity, their role, and the task at hand.

Datahaven’s tightly integrated collection of SDKs, APIs, visual components, and related ECM tools allow solutions providers to create robust, content-enabled versions of these custom, data-centric applications in as little as one day. Our products can be embedded directly into the already-familiar user interface of an existing application—whether built on a client-server, web-based, or legacy architecture. The results can be compelling.

Implemented through a series of integrated visual components, Datahaven not only drastically shortens the time-consuming software development process associated with ECM functionality, it also provides the most seamless user experience possible.

Embedding a Scalable, Enterprise-Class Content Repository

Datahaven offers complete, programmatic control over its advanced functionality, including access to an open, flexible, and scalable storage repository. The Datahaven repository

- stores scanned or faxed images at the page-level
- allows the import of content objects of any type (e.g., emails, audio, and other electronic files)
- gives solutions providers the flexibility of defining the storage media as a default setting across all content types or defining the storage media by the type of content object to be stored
- offers migration tools and functionality to help move content from one storage medium to another
- enables searching of any metadata associated with the content object

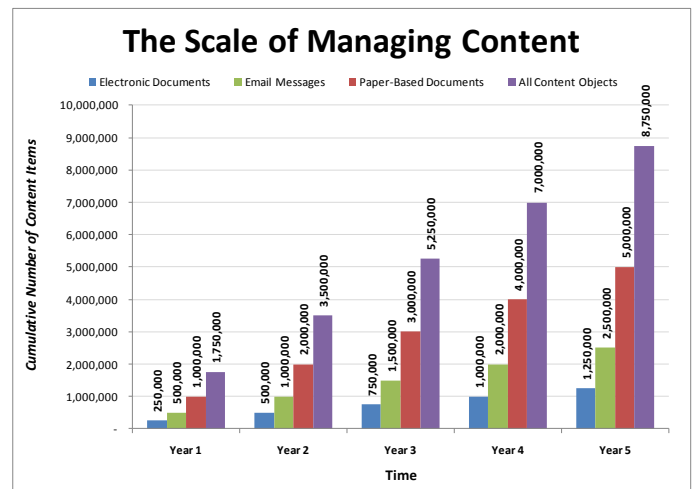


Figure 6: The scale of managing content in an organization of 1,000 information workers

The flexibility and scalability of the underlying repository of a content-enabled solution is crucial. Datahaven supports multiple database platforms, including SQL Server, SQL Server Express, Oracle, and Informix. Datahaven also supports the use of a wide array of other storage technologies, including file-based and

UNC-addressable file systems. With current implementations (some of which have been in production for more than 10 years) managing content objects well into the tens of millions of pages and hundreds of gigabytes, the Datahaven repository meets the challenge.

If we re-visit our hypothetical organization or department of 50 employees each of whom works with an average of 250 electronic documents, 500 emails, and 1,000 paper-based documents, we can see that the number of items stored in the repository will grow by 1.75 million items each year assuming no growth. In five years, the repository will have to manage nearly nine million items of content.

Embedding a Context-Sensitive Virtual File Cabinet

One of the most unique and valuable elements of a Datahaven solution is its virtual file cabinet. It can be embedded practically anywhere within the existing user interface of a data-centric application. At its most basic level, the Datahaven virtual file cabinet affords your application the ability to display a view of items of content in a list or grouped within a parent-child folder structure. Thus, on the face of it, the Datahaven cabinet appears simplistic. After all, its interface looks just like the familiar parent-child folder structure of a shared network drive.

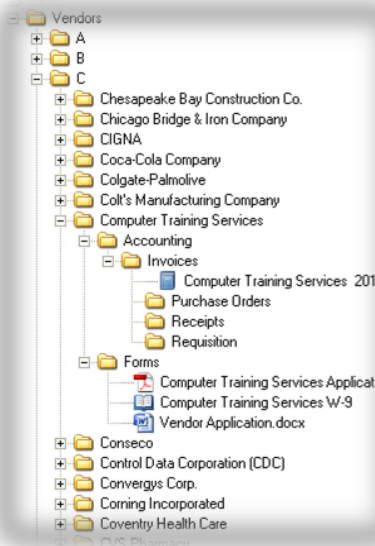


Figure 7(a): The larger Datahaven content repository for our sample solution . . .

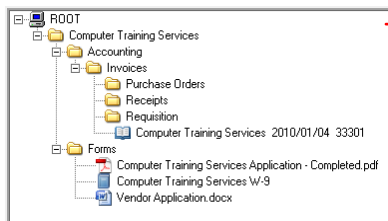


Figure 7(b): ... is programmatically-filtered based on the user's identity, role, and application context so that . . .

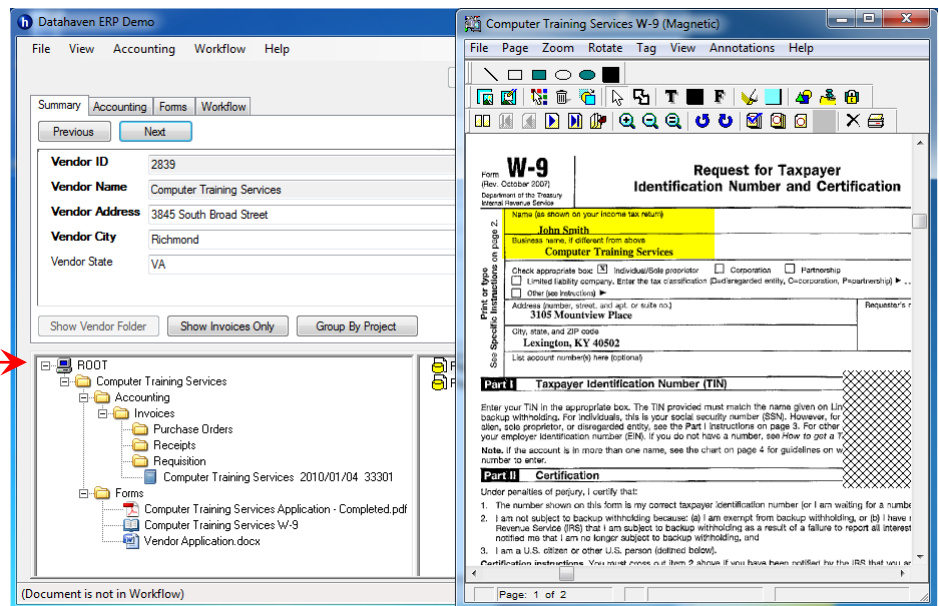


Figure 7(c): ... any screen of the content-enabled solution displays only the content relevant to that user in a particular circumstance.

Don't let the simplicity of its interface fool you. The Datahaven virtual file cabinet delivers much more than a shared network drive.

Once embedded within your data-centric application, the Datahaven cabinet offers a powerful, enterprise-class tool that greatly expands the functionality and value of a data-centric application. For example, it provides users with familiar right-click access to a wide range of sophisticated document and content management features such as splitting and merging documents, workflow, and adding pages to existing documents. It also provides your otherwise data-centric application with a consistent, intuitive extension of its interface to provide authorized users with instant, security-controlled access to relevant content of any type stored in the enterprise's larger content repository.

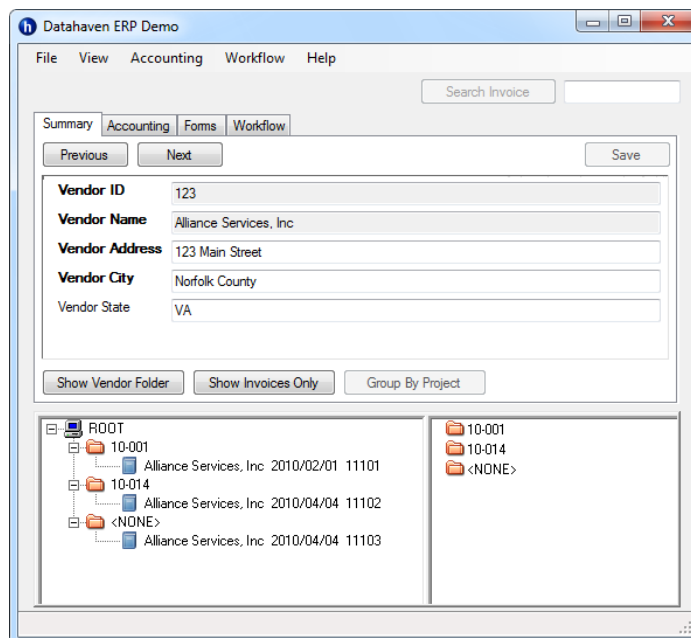


Figure 8: Example of a dynamically-changed view of the Datahaven virtual file cabinet created by programmatically filtering and grouping content based on its metadata

The Datahaven virtual file cabinet provides infinite flexibility over how your data-centric application displays content and can even allow users to change the “default” view to one more helpful to a unique task. A user can view content by modifying and grouping folder structures on-the-fly. For example, the content items stored in the Datahaven content repository can be generated, filtered, re-arranged, and displayed to each end-user based on any combination of variables, including

- the user's identity and security permissions
- the record displayed on a particular screen
- the active tab of the data-centric application and
- the metadata associated with the underlying record or its related documents.

This allows any screen of your data-centric application to display related content with little to no input needed from the user. Finally, your application can include buttons and other controls that change the default folder structure on a pre-defined basis.

Leveraging OCR & Other Automated Data Extraction

Using automated data entry technologies like barcode recognition, OCR, OMR, and ICR has been shown to increase administrative productivity by as much as three-to-five fold when compared to the use of manual data entry processes. For example, a zonal OCR process can automatically extract the metadata required to add a record to the data-centric application with no additional human involvement. In addition, Datahaven could use this same metadata to apply an automated naming and filing convention so that every document is named and filed the same way every time.

This automated naming and filing process alone can dramatically reduce human error and administrative overhead costs. In combination with Datahaven's workflow and audit trail features, our content-enabled solution can greatly increase the organization's ability to satisfy its legal and regulatory recordkeeping requirements.

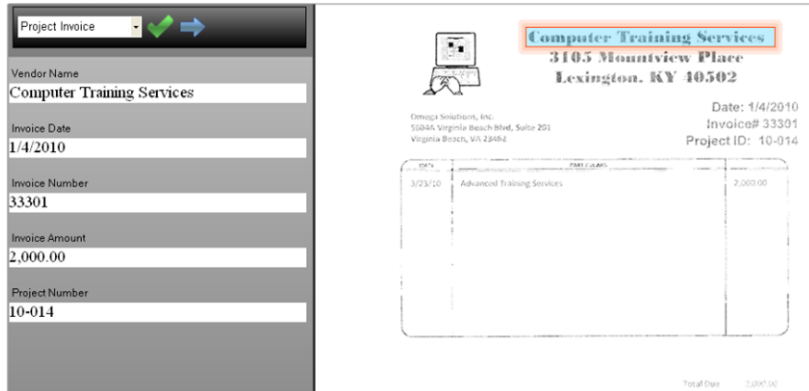


Figure 9: An example of the ability to use advanced OCR technology to automatically extract data from a document

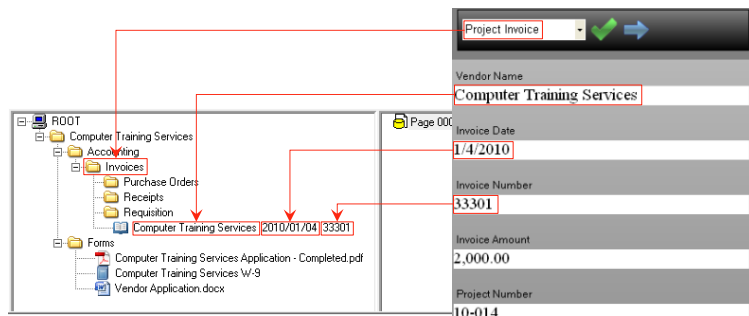


Figure 10: An example of the use of automatically extracted metadata to name documents and programmatically file them in the correct folder structure for this vendor.

Embedding Content-Centric Workflow Components

Your data-centric application can embed Datahaven's integrated, document-centric workflow component in a variety of ways, including within a portion of a screen, on a dedicated tab, or as a pop-up window. In the sample application in Fig. 11, we simply embedded the Datahaven workflow component on a dedicated tab. Once embedded, Datahaven allows the data-centric application to programmatically support end-users' custom workflow needs, including pre-defined and ad hoc workflow processes.

From here and in other areas of this application, Datahaven allows users to process documents and other content literally within minutes of its receipt. For example, based on their role and their security permissions,

workers can

- access the same workflow queue simultaneously,
- have access only to a single content item based on business rules to avoid cherry-picking,
- determine whether to approve the item or route it for others to review,
- add annotations and user notes,
- split and merge documents, and
- monitor process bottlenecks in real-time and re-assign work as necessary.

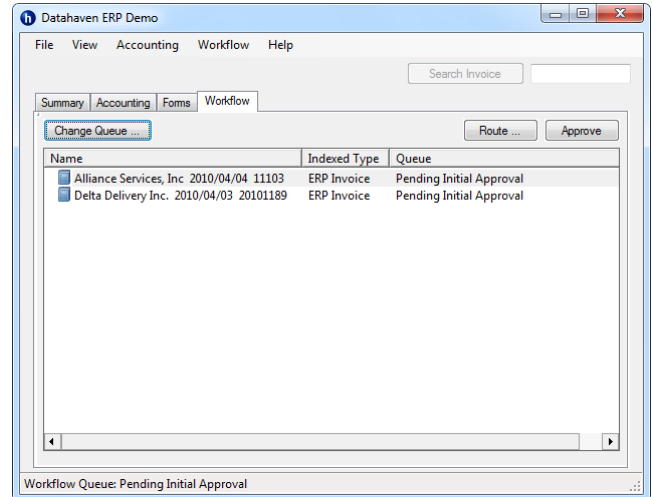


Figure 11: A list view of content in a sample workflow queue

Accessing Content (Not Searching For It)

Nearly all stand-alone ECM solutions on the market limit a user's ability to access content within their repository to a search screen and/or a workflow queue. This limits a user's access to those content objects listed in a specific, pre-existing queue at a particular moment in time or to those content objects produced by search terms (that hopefully are broad enough and accurate enough).

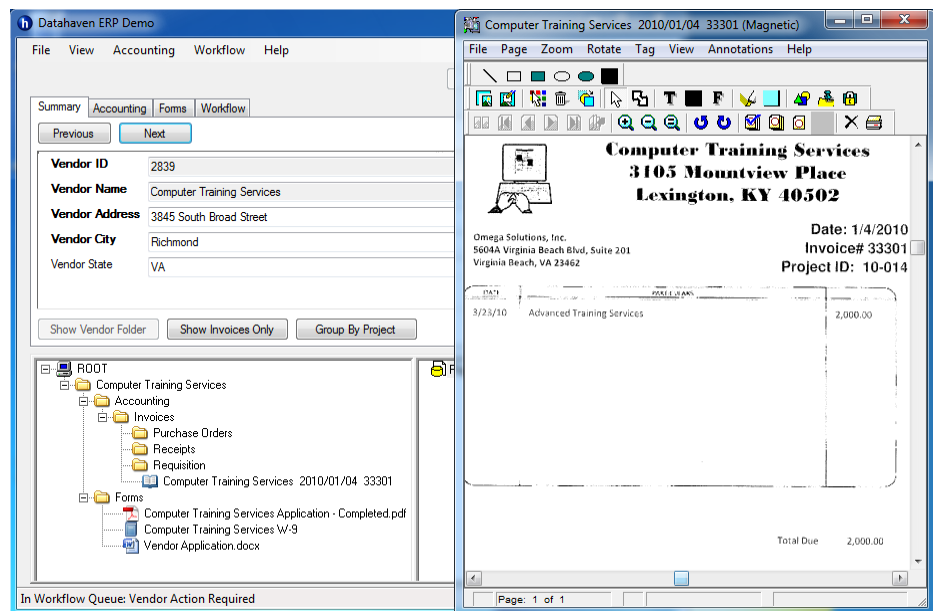


Figure 12: An example of how an information worker has instant access to any content related to this vendor from a single screen in our sample ERP application.

With Datahaven, all relevant content simply appears where it's needed based on the user's identity, role, and the task at hand. Accessing it within the Datahaven virtual file cabinet only requires the information worker to click. Of course, information workers also can find content via a search of its metadata and in workflow queues as well.

Securing Access to Content & Auditing User Activities

All of this functionality would be of limited value if we could not secure who has access to it and to the content itself. Datahaven offers granular security settings by enabling your application to apply privileges and permissions to actions, folders, and/or documents at the user and/or group levels. You can choose from at least three options when integrating Datahaven-related security into your application.

- First, you can expand your application's existing security model for data and set Datahaven's privileges and permissions so they can be applied to users and groups whenever they access content stored in the Datahaven repository.
- Second, it can choose to use Datahaven's security model and pass the user or user's group membership programmatically to Datahaven thereby allowing the user's privileges and permissions to be set automatically as set up in the Datahaven security model.
- Third, your application can leverage Active Directory.

Additional page-level security can be enforced on a particular area of a document by applying security redactions such that authorized users can see through the redaction, but prevent unauthorized users from seeing the information hidden by the redaction. This creates a significantly higher level of security for sensitive documents than could be achieved using file attachments, file cabinets, or personal email archives.

Audit trails are generated based on user actions and server-based activity. Audit trails in Datahaven can track and record every action at the document level, including the date, time, user, and action performed. Managers and executives can use the granular information tracked in the audit trails maintained by Datahaven to generate a wide range of reports.

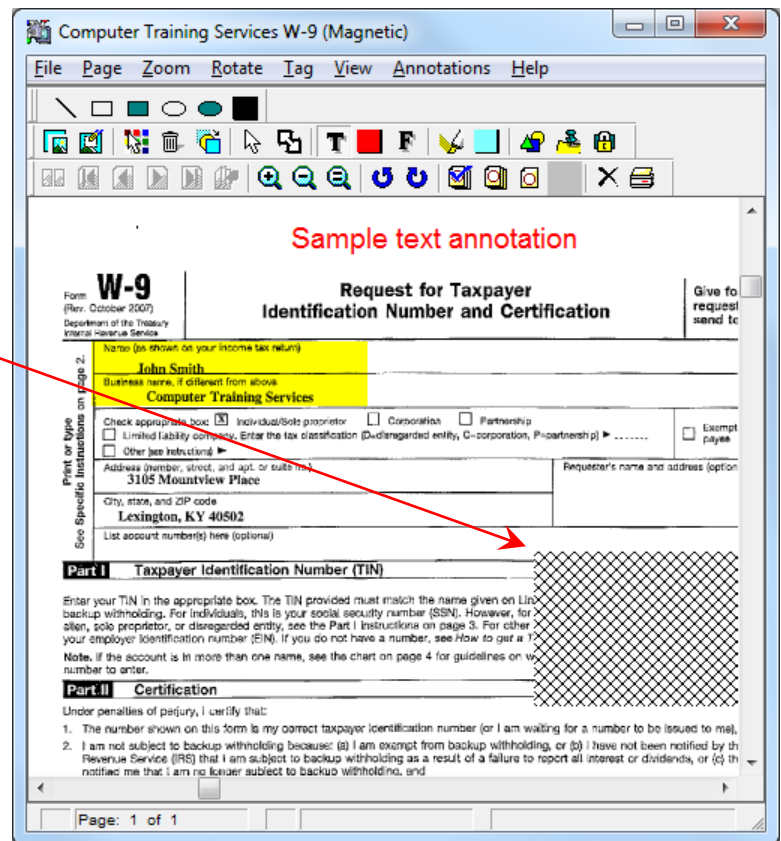
The image shows a screenshot of a PDF document titled "Computer Training Services W-9 (Magnetic)" displayed in a software application window. The document is a "Request for Taxpayer Identification Number and Certification" form. A red arrow points to a large, cross-hatched redaction box that covers the "Part I Taxpayer Identification Number (TIN)" field and the "Part II Certification" section. The form contains the following visible information: Name: John Smith; Business name: Computer Training Services; Address: 3105 Mountview Place, Lexington, KY 40502. The form also includes checkboxes for "Check appropriate box" (Individual/Sole proprietor, Corporation, Partnership, Limited liability company, Other tax treatment) and "Exempt payee".

Figure 1: An example of a security redaction that prevents unauthorized users from viewing sensitive information

CONCLUSION

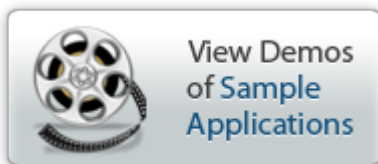
This white paper defined the concept of data-centric software applications. It identified the pervasive problems (and opportunities) related to information workers' frustrations when attempting to find paper, faxes, emails, and other content related to records managed by these one-dimensional applications. It also described the shortcomings of trying to solve these problems with the traditional approaches afforded by file attachments, low-level SDKs, and stand-alone ECM systems.

The Datahaven SDK for ECM offers a unique alternative that allows solutions providers to embed components for document and content management directly into an existing UI of custom-written and popular third-party applications. Our unique, tightly-integrated development solutions for ECM allow solutions providers to embed directly within the user interface of existing software applications—whether custom-written or one of the more popular third party enterprise applications. Datahaven's collection of SDKs, APIs, visual components, and other ECM tools transform data-centric applications into what analysts have referred to as "content-enabled," "content-rich," or "content-centric" solutions.

From document imaging to email and fax integration, OCR data capture, workflow, and content publishing and distribution, Datahaven provides the means to enable a data-centric application to natively capture, process, store, and manage the entire spectrum of content and related metadata from within its already-familiar user interface. This provides information workers access—from directly within the applications they rely on every day to do their job—to precisely the right content at the precisely the right time based on their identity, security permissions, and the context of their task.

With a scalable, enterprise-class repository and a context-sensitive virtual file cabinet, solutions providers have the tools they need to add advanced ECM features directly into the enterprise applications that information workers use every day. The content-enabled solutions that result will provide information workers with instant access—from an already-familiar UI—to the complete record, including all related paper, faxes, and email, within a single screen.

We cannot cover every feature or possible use case scenario in a white paper. We have posted recorded demos of additional examples of content-enabled software solutions on our web site. You can also contact us at any time to arrange for a technical review of your opportunity or schedule a live demo of our technology.



*Contact us today so we can show you
how to content-enable an existing
application in as little as one day.*



Torrential Data Solutions offers complete, tightly integrated, database-driven software development tools for ECM—enterprise content management—that corporate developers, systems integrators, and ISVs can embed directly within the UI of existing software applications. Our collection of SDKs, APIs, visual components, and other tools transform these data-centric applications into content-enabled solutions in as little as a day. From document imaging to email and fax integration, OCR data capture, workflow, and content publishing and distribution, Datahaven provides the means for these data-centric applications to natively capture, process, store, and manage the entire spectrum of content and related metadata from within its already-familiar user interface.

