Building a Consumer Interface for Documentum

Case Study

October 2010

www.tsgrp.com

Introduction

Purpose

This case study details Technology Services Group's experience building simplified interfaces for consumers of content stored in Documentum.

Business Issue

While a typical Documentum application (ex: Webtop) provides a "one stop shop" for authors and approvers, the interface can be challenging when "consumers" are just looking for quick search and retrieval. The consumer users are frequently overwhelmed with the full-featured interface that is necessary for authors and sometimes approvers. Businesses have approached TSG or other consulting firms to address basic concerns with a one-size-fits-all approach. We repeatedly see the following two requirements for a consumer interface:

- Simplified Interface – As presented later in this paper, the interface should be simple and not require training to deploy to a large group of consumers.
- Performance – Should be quick in regards to response time and navigation as consumers can often be frustrated by performance delays.

Other benefits of a consumer interface for Documentum could include:

- Business Continuity – To provide availability to documents when Documentum or other system components are down for maintenance, upgrade, or migration
- Independence – From document management releases and the ability to view more than just Documentum documents.
- Reduced Load – Both in terms of processing and licensing on the main Documentum server.

Example Case Studies

This section will present three representative examples of clients that have implemented consumer interface solutions. In regards to TSG experience, the consumer interface is the most popular and successful best practice for our Documentum clients. TSG has implemented this solution across multiple industries for multiple clients. Below are three representative case studies for reference.

Large Life Sciences Company

Prior to 2005, TSG had developed a number of consumer interfaces as either small additions to Webtop or complete standalone interfaces that accessed Documentum real-time. In 2005, a large life sciences company hired TSG for a substantial upgrade and implementation of a Documentum system built for access to controlled documentation. At the time, the client was planning on a migration effort that was estimated to take two weeks. During the downtime, the client wanted a small, temporary, system to access just the approved content out of Documentum. TSG constructed a cached repository of documents and developed a simple interface to gain access to the content.

During the two weeks of downtime, the consumer interface was a huge success. Users liked the ease of use as well as the higher speed compared with Documentum Webtop. During the migration, it became obvious to the client that the system would not be successful unless the stand-alone interface was a permanent part of the ongoing production system. From this experience, TSG began to formally recommend clients to consider a consumer interface as part of their production system.

Some relevant points in regards to the implementation:

- OpenMigrate rather than Site Caching Services – In production, the solution required constant monitoring of the docbase looking for new and approved content. The interval to look for new content was every three minutes. Site Caching Services was determined to have too much overhead for this solution and the decision was made to leverage OpenMigrate instead.
- Database Approach – Consistent with Site Caching Services, OpenMigrate pushed content to a file store and database to allow for attribute searching.
- PDF Aqua Integration – Approved documents required watermarks with date-time stamps. The interface leveraged PDF Aqua for this capability consistent with the Webtop application.
- OpenOverlay – The initial system only used PDF Aqua to add overlays at publish time. Once the permanent solution was put in place, our client wanted to add a dynamic overlay at view time listing the user ID and date the document was viewed/printed. This dynamic overlay could not be added with PDF Aqua since we wanted the system to remain completely separate from Documentum for business continuity purposes. As a result, OpenOverlay was used to satisfy the addition of the dynamic overlay at print time.

International Shipping Company

For this company, access to Documentum was a priority but the need to leverage a cached repository was not required. TSG developed an updated search using our High Performance Interface (HPI) product based on the needs of the client as well as the lessons learned from previous consumer interface installations. Unique components for this client that were later incorporated into our High Performance Interface open source offering included:

- OpenContent Integration – OpenContent provides an open source web service layer capable of accessing Documentum, Alfresco, SharePoint or even the file system for our cached clients. By developing on OpenContent, TSG was able to provide the company (on Documentum) a quick and easy solution that could support both Documentum and non-Documentum repositories. As the product was enhanced for other non-Documentum customers, the company was able to leverage the same product updates on Documentum.
- Export to Excel – provided for some reporting capabilities and allowed for exporting of search results to Excel.
- Saved Searches – Provided for the ability to save searches per user for quick access to information that is regularly queried.
- Google Web Toolkit – provided a more enhanced user interface that performed consistently across varying internet browsers.

Large Life Sciences Company

In the second half of 2010, TSG worked with one of our oldest clients for their first implementation of a cached consumer interface, the consumer interface was used to provide access to their controlled manufacturing documents. The primary driver for this application was business continuity although the application did satisfy other needs including interface design, speed and license issues. While leveraging HPI and OpenMigrate, some unique requirements also included:

- The ability to conduct both full-text and attribute searches
- No backend relational database required for implementation due to license costs
- The ability to integrate into other enterprise search tools from other platforms.

For this client, Lucene was identified as the full-text indexing tool as well as the database repository. Leveraging Lucene for both provided the ability to combine full-text and attribute searching while also not requiring another Oracle license. The majority of the software infrastructure (Lucene, Linux) did not require additional purchases.

The solution conformed to OpenSearch standards. (http://www.opensearch.org), OpenSearch standards are a collection of simple formats for sharing of search results. The solution allowed for a separate application (SharePoint in this case) to execute a Lucene search and return it in an OpenSearch compliant RSS feed. The solution allowed SharePoint users to access content stored in Documentum without individual license

concerns and Documentum user-ids (and associated licensing) while maintaining the integrity of the Documentum security.

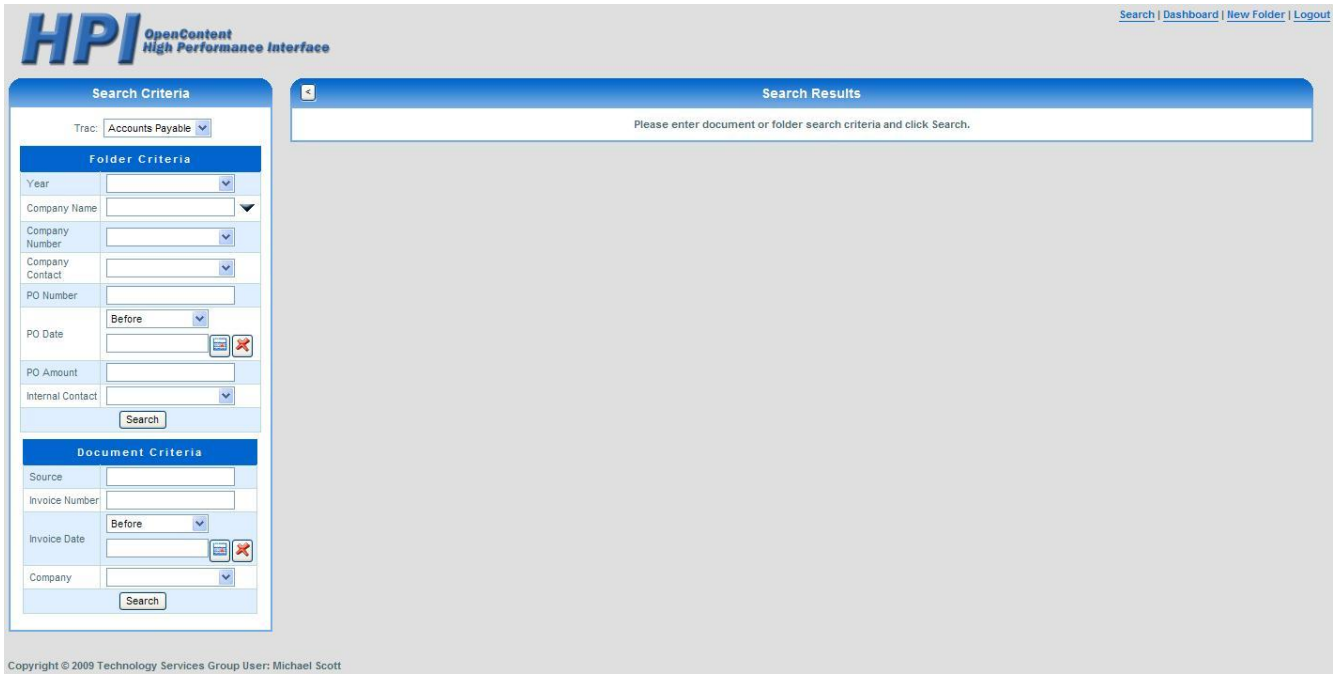
Webtop vs. HPI Search

This section of the whitepaper will illustrate differences between Webtop to a consumer interface based on best practices. HPI Search reflects all of the interface design lessons learned from TSG.

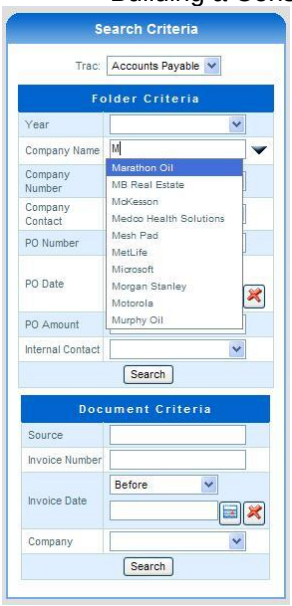
HPI Search & Retrieval

The HPI Search and Retrieval application consists of a Search screen and a Document Details screen.

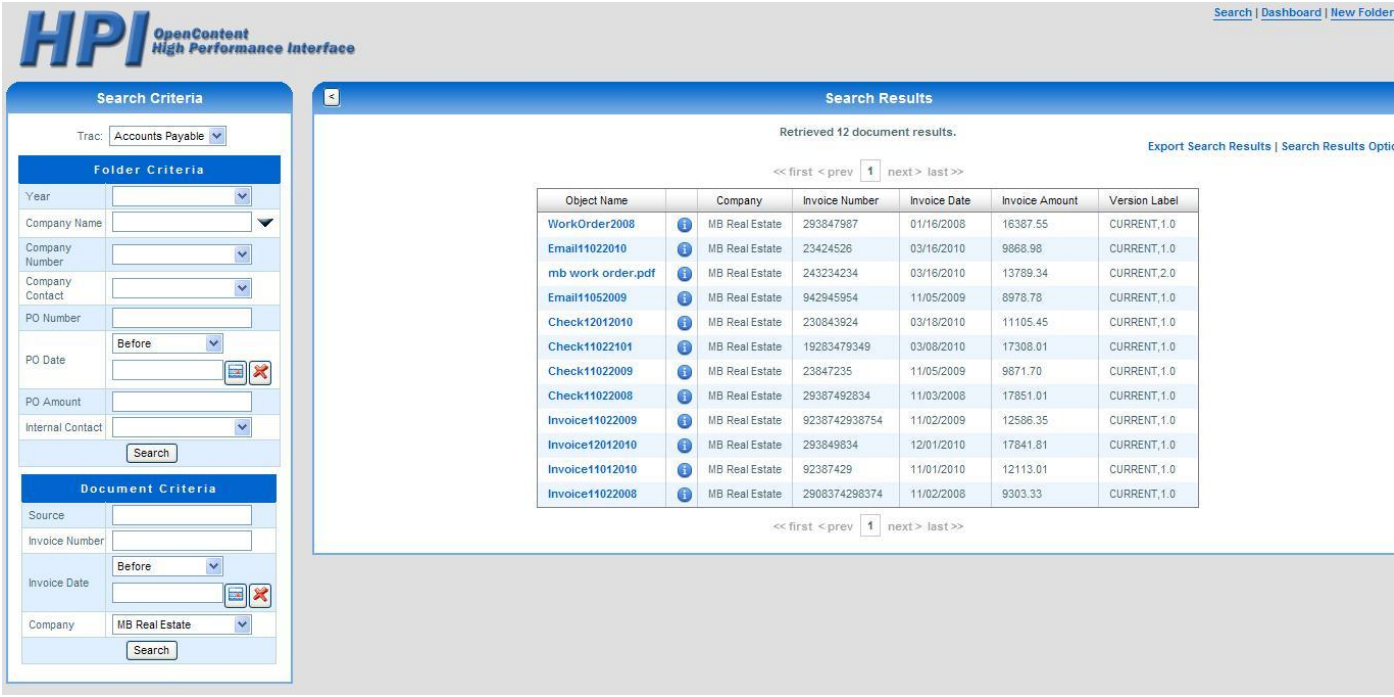
- Ssearches can be quickly executed from one page.
- The search Criteria and Search results are displayed on one page.
- Display Options can be configured on the fly and are saved for each individual user.
- Sorting and Pagination can be performed without re-executing the search.
- Document Properties can be viewed at any time without navigating away from the search results.



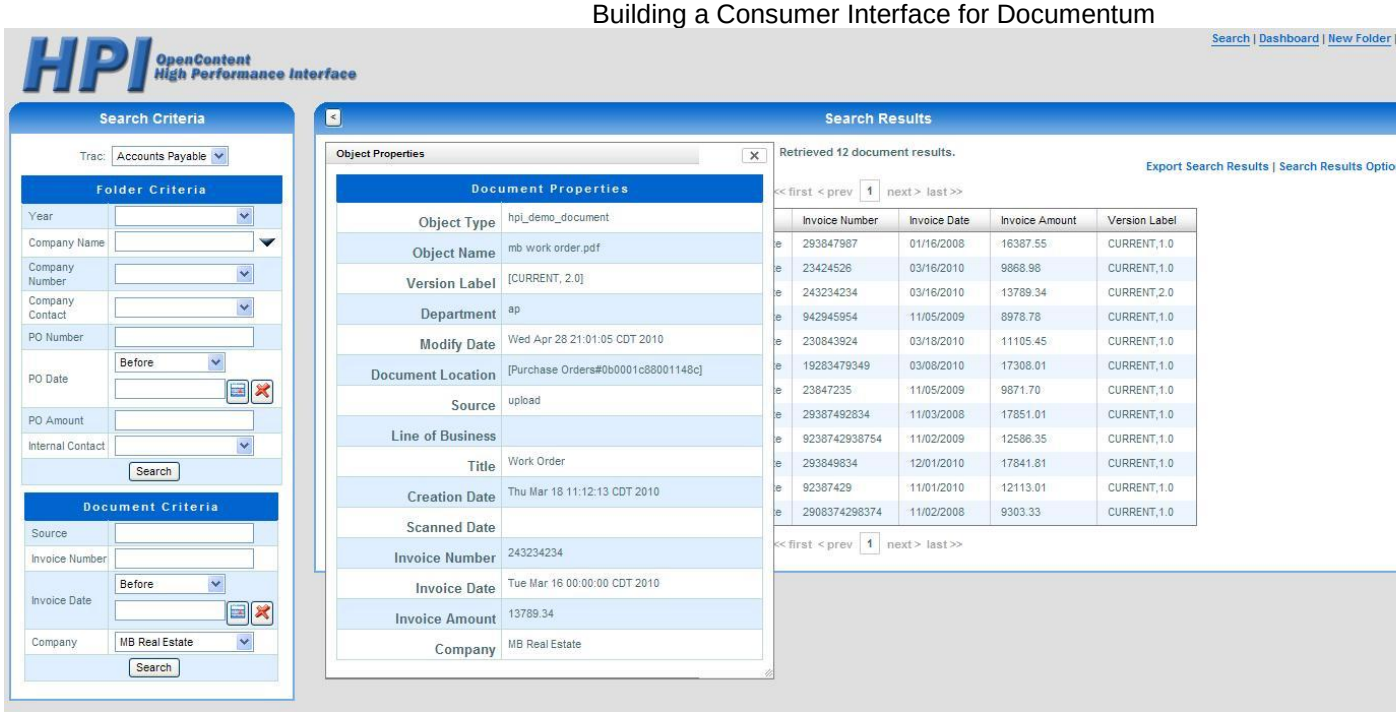
HPI Search Screen prior to a search.



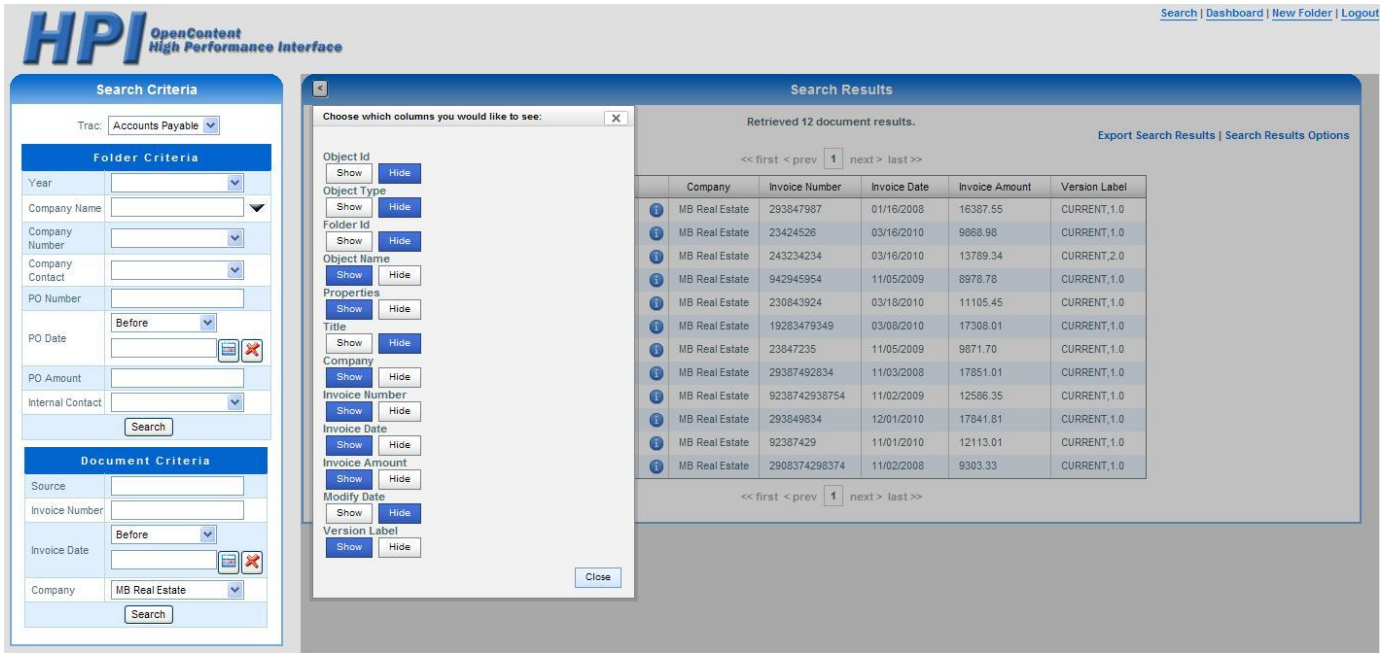
Entering text into an HPI type-ahead field for search.



HPI Search Screen search results.



HPI Search Screen Document Properties.



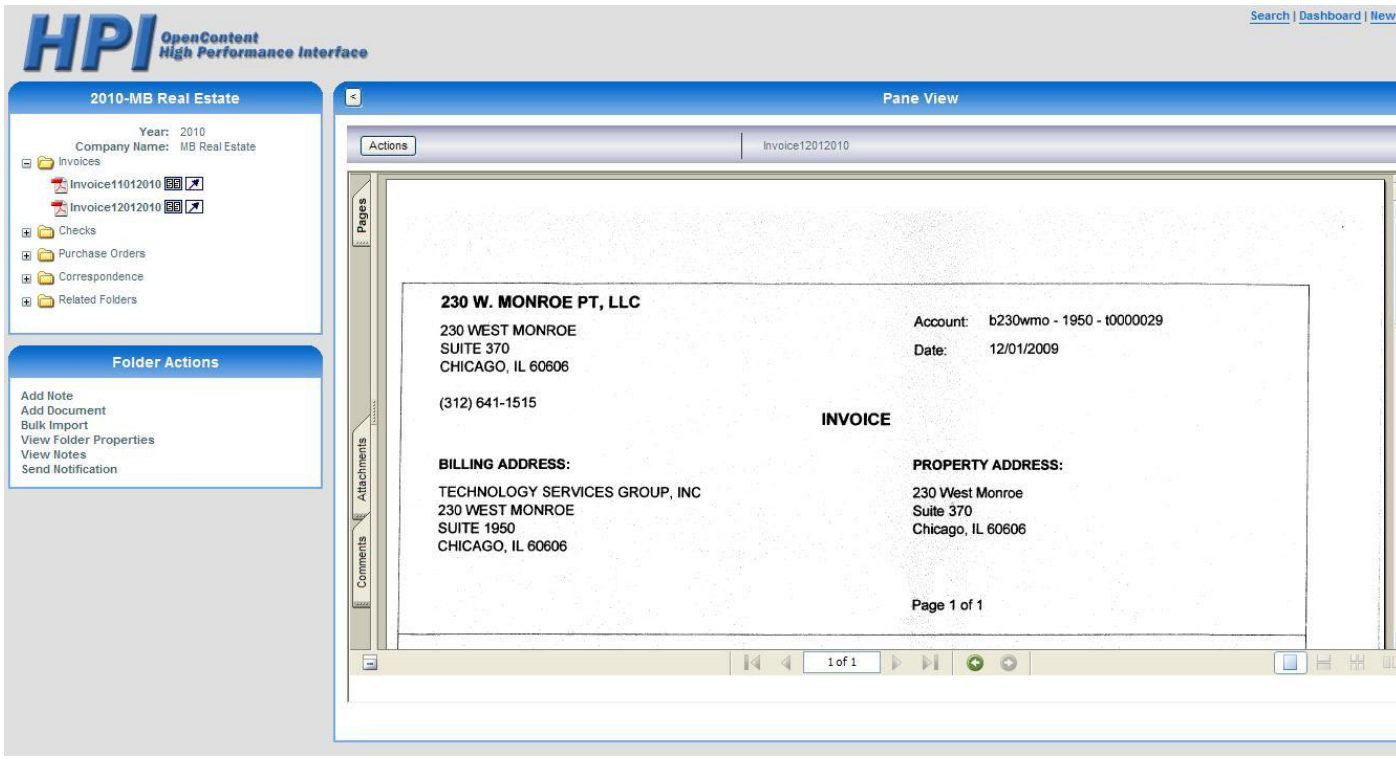
Configuring search columns for viewing 'on the fly'

<< first < prev 1 next > last >>

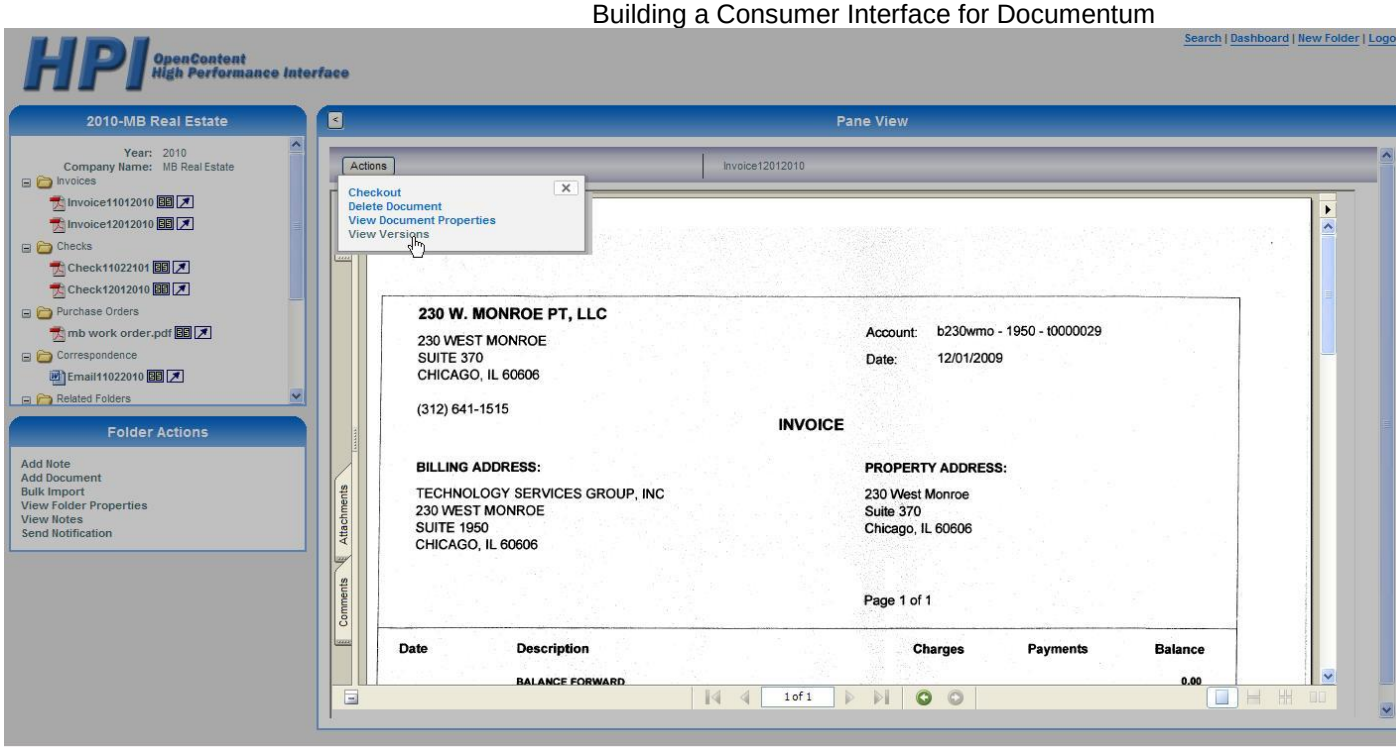
Object Name		Company	Invoice Number	Invoice Date	Invoice Amount ▲	Version Label
Check12012010		MB Real Estate	230843924	03/18/2010	11105.45	CURRENT,1.0
Invoice11012010		MB Real Estate	92387429	11/01/2010	12113.01	CURRENT,1.0
Invoice11022009		MB Real Estate	9238742938754	11/02/2009	12586.35	CURRENT,1.0
mb work order.pdf		MB Real Estate	243234234	03/16/2010	13789.34	CURRENT,2.0
WorkOrder2008		MB Real Estate	293847987	01/16/2008	16387.55	CURRENT,1.0
Check11022101		MB Real Estate	19283479349	03/08/2010	17308.01	CURRENT,1.0
Invoice12012010		MB Real Estate	293849834	12/01/2010	17841.81	CURRENT,1.0
Check11022008		MB Real Estate	29387492834	11/03/2008	17851.01	CURRENT,1.0
Email11052009		MB Real Estate	942945954	11/05/2009	8978.78	CURRENT,1.0
Invoice11022008		MB Real Estate	2908374298374	11/02/2008	9303.33	CURRENT,1.0
Email11022010		MB Real Estate	23424526	03/16/2010	9868.98	CURRENT,1.0
Check11022009		MB Real Estate	23847235	11/05/2009	9871.70	CURRENT,1.0

<< first < prev 1 next > last >>

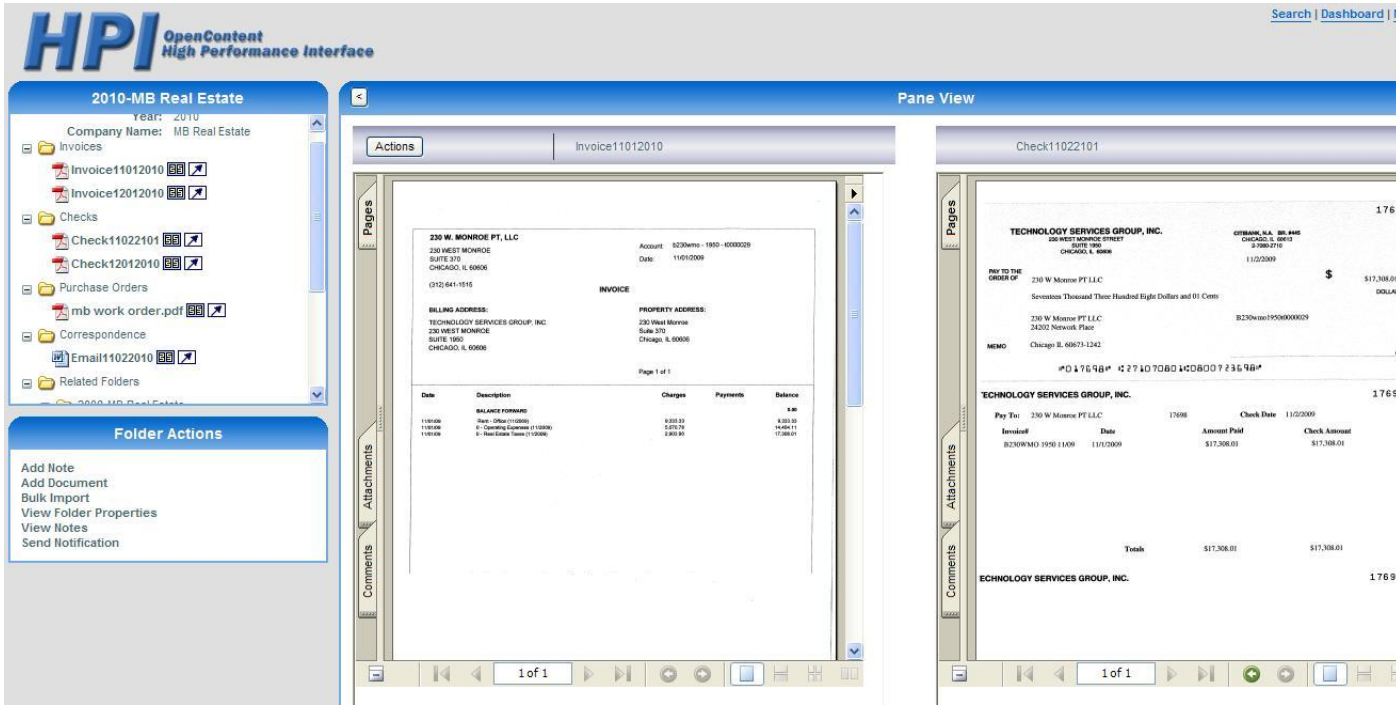
Sorting search results



HPI Search Document View.



Selecting a document action – View Versions.

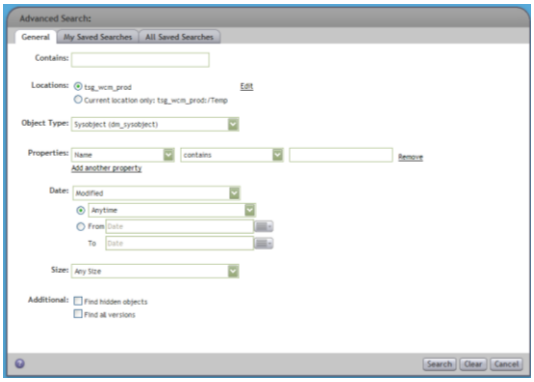


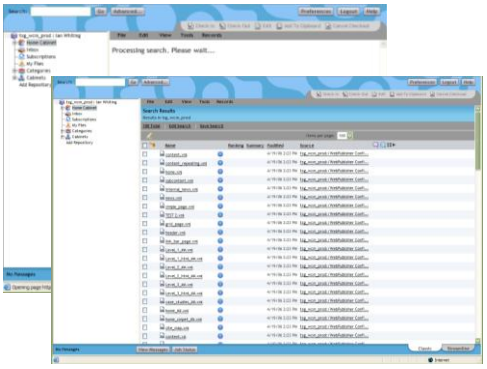
Document dual-pane view.

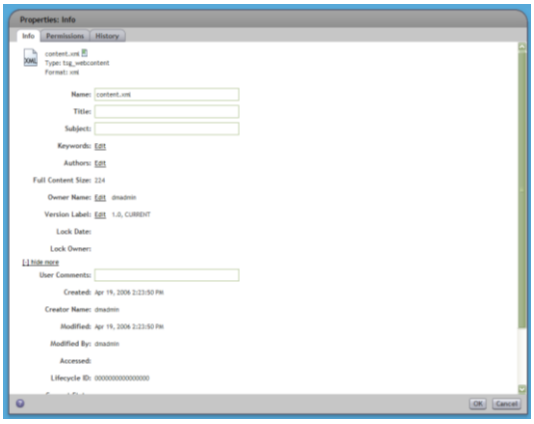
Documentum Webtop Comparison

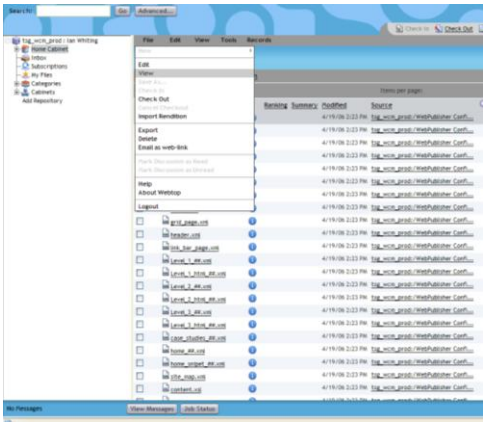
Within Documentum Webtop, searching is built as a "one size fits all" application. Through a generic interface, users are walked through the process of developing a query by choosing document types, attributes, expressions and values. The HPI Search and Retrieval interface simplifies the process by presenting default document types and expressions.

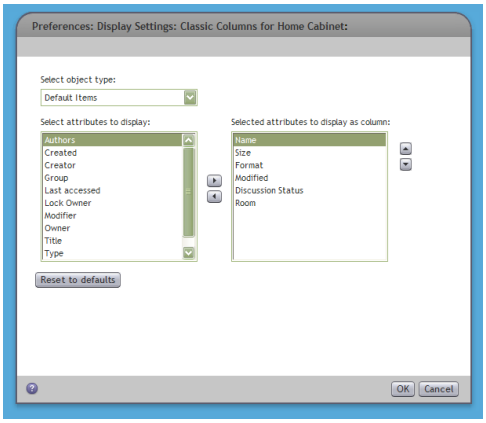
For reference, the Documentum Webtop screens are presented below.

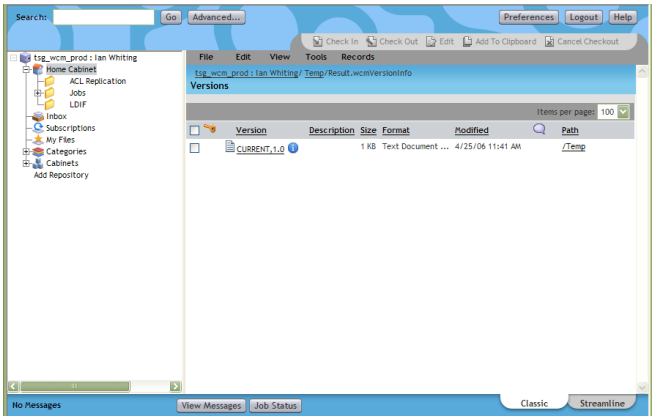












Specific limitations of Documentum Webtop search compared to the HPI Search and Retrieval application include:

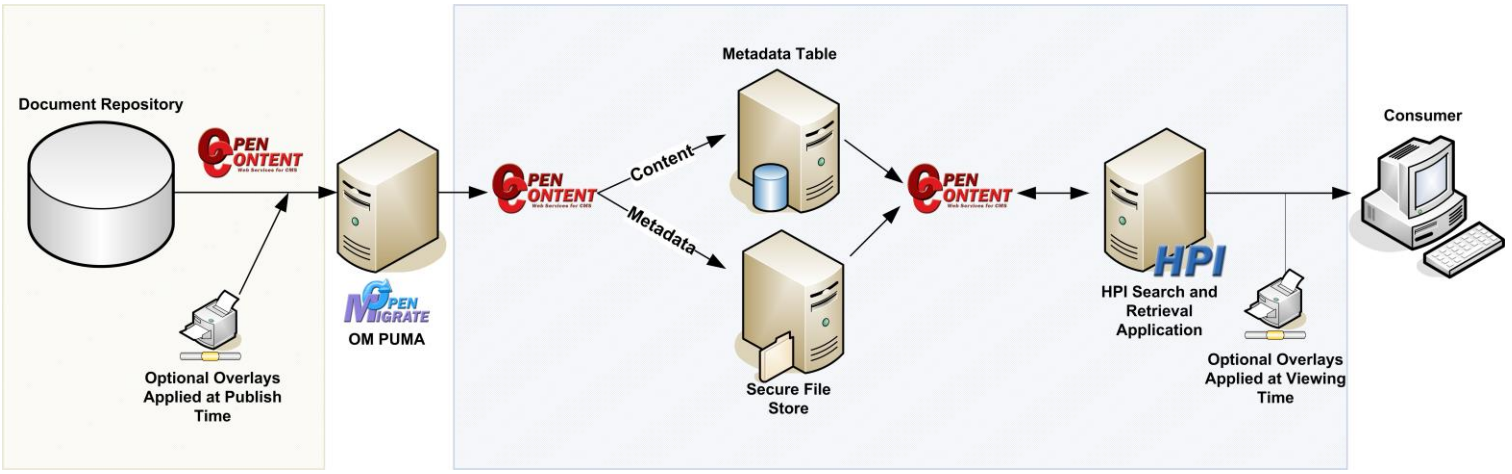
- In Webtop, Advanced Searches must be done from an Advanced Search page.
- Webtop does not provide as many options for search criteria fields (i.e. no type-ahead, no value assistance drop downs)
- Webtop cannot display search criteria and results on same screen
- Webtop display options cannot be configured on the fly. Every time a user performs a search they must redefine which attributes are displayed.
- Webtop sorting and pagination cannot be performed without reloading the entire page.
- Webtop document content is displayed outside of the browser.
- Webtop document versions are only displayed from another page.

Cached Repository Approach

This section of the document will present specific details on the Cached Repository approach. As mentioned previously, HPI (based on OpenContent) can provide solutions for both Documentum real-time searching as well as cached repository searching.

Cached Repository Overview

OpenMigrate and HPI have been used by large companies as a way to provide business users with a faster and secure way of accessing documents outside of Documentum.



As depicted above, OpenMigrate is used to synchronize content from Documentum to a metadata table (database) as well as a secure file store. OpenMigrate can be configured to move content once it reaches a certain stage (ex: Released) as well as remove content when it has changed stages (ex: Superceded). OpenMigrate can run at a set interval defined by the business (e.g. every three minutes)

When the data is extracted, a PDF overlay (i.e. PDFAqua or OpenOverlay) can be applied, if required.

Typically the metadata table will be a standard database or Lucene as pointed out in one of the earlier case studies. In this manner, HPI can access the released content outside of Documentum for higher performance and reduced Documentum licenses. In the event the Production Documentum Repository experiences downtime (upgrade, maintenance), users will still have access to the vital production documents.

OpenMigrate can be customized to only migrate certain document types, statuses, versions, or any combination of criteria. Other components of the cached approach include:

- **User Security** - The business can define user roles and access levels depending on their needs. Users have permissions that can be setup to mimic ACLs in

Documentum. Administrators have the ability to add users, delete users and even change a user's permissions. New users can be given access to HPI without having to purchase additional licenses from your ECM provider. This is ideal for users who only use their ECM to search and view content.

- **Minimize Upgrade Impact** – Since the cached repository exists outside of Documentum, users would continue to have access to the application and documents during any Documentum upgrade (ex: 5.3 to 6.X). Once the upgrade is complete, OpenMigrate can be easily re-pointed to begin synchronizing content from the new repository eliminating consumer downtime.

Advantages of a Cached approach

Typical ECM database architecture is created from a document management perspective. Every document type is a series of multiple tables all joined together by a common object identifier. This design is not conducive to quick search and retrievals.

The OpenMigrate PUMA flattens these very complex tables into a single table similar to those familiar with Documentum Site Caching Services. The HPI Search and Retrieval Application can then execute a single SQL statement against one simple database table.

Documentum Webtop Performance Comparison

Performance of the HPI Search and Retrieval Application when compared to Documentum Webtop can be divided into two distinct areas.

1. **System Performance** – Retrieval time (time from completing the search to seeing results) is greatly improved. This primarily due to the simplified database architecture within the cached repository. These results will vary depending on document taxonomy and specific ACL configuration between Documentum installations.

   Data – Simple Query – 10,000 entries
   - Webtop: 15-60 seconds
   - HPI Search and Retrieval: sub-second

2. **User Performance** – Another evaluation point should include user statistics for how quickly the user can enter the query itself. Given the complicated nature of Documentum, the users typically spend at least two minutes configuring (or reconfiguring) the search. With HPI, that timeframe is reduced to seconds.

Conclusion

As presented throughout this paper, building a consumer interface can provide some clear benefits for Documentum users in regards to:

- User Acceptance
- User Performance
- Reduced Training and Support

If a "cached approach is used", additional benefits can include:

- Business Continuity
- Better performance
- Reduced load in regards to performance on the main Documentum system and licensing
- Ability to "cache" non-Documentum content for access within the same search/retrieval interface.

The High Performance Interface (HPI) from TSG (certified by Documentum) represents many of the "best practices" for developing a consumer interface. Additional information including screencams and source code is available on our Website at:

http://www.tsgrp.com/Open_Source/OpenContent/open-content-hpi.jsp

OpenMigrate from TSG (certified by Documentum) is commonly used as a component of a cached approach. Additional information and source code is available on our Website at:

http://www.tsgrp.com/Open_Source/OpenMigrate/open-migrate.jsp