



Empowering  
Visual Search

# Documill Visual Search Suite

2011-10-13

*Disclaimer: All information included in this presentation is company confidential, and should not be redistributed in any media or form without prior written permission from authorized Documill personnel. © 1996–2011 Documill. All rights reserved.*

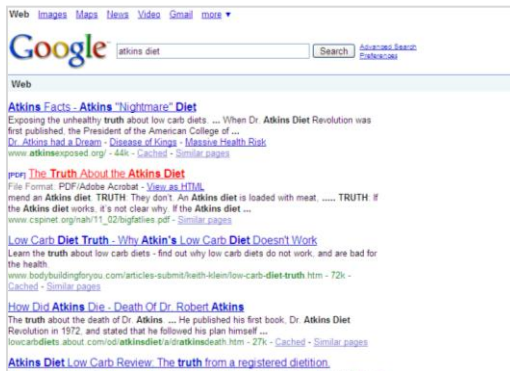
# Documill Offering

- Helps You to Leverage Your Existing ES/ECM Investment !
  - Deployed on top of existing content indexing and search capabilities
  - Great fit for any system dealing with documents; document libraries, search solutions and content management/distribution
  - Enables previewing, re-using and sharing of enterprise documents
- Documill Visual Search Suite - Applications
  - Documill Resultset Previewer ("The Ribbon")
  - Documill Online Document Viewer
  - Documill Floating Viewer
  - Documill Page Collections ("Visual Bookmarking")

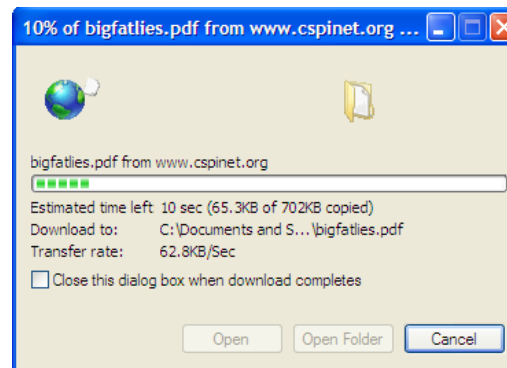
# Searching Enterprise Documents is Frustrating

## PDF, DOC/X, PPT/X...

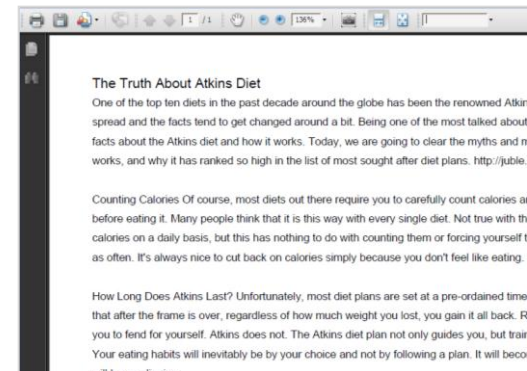
Search



Download



Search again



**...because enterprise documents are different (than HTML):**

- Often very long documents
- Include important visual objects, which are ignored by text based indexing
- Just a few inbound links, practically no outbound links
- When the right document is found, getting to the right page requires additional search.

# Visual Search for Documents

## Normal search result list

Home > Search Results

All Sites

documill

Preferences Advanced

Result Type: 1-9 of 9 results

Any Result Type

Adobe PDF

Site: Any Site

www.documill.com/en

Author: Any Author

Mika Könölä

Mika Könölä, Ter...

Modified Date: Any Modified Date

Earlier

**Documill Dynamo**  
Interactive customer communication (c) 2009 Documill, Inc. All rights reserved.9About Documill:  
**Documill** is an independent software vendor (ISV ...  
Authors: Mika Könölä Date: 11/23/2009 Size: 1MB  
[http://www.documill.com/en/downloads/20091123\\_Documill\\_Dynamo.pdf](http://www.documill.com/en/downloads/20091123_Documill_Dynamo.pdf)

**Enhancing e-Discovery Experience with Documill Visual Search**  
Enhancing e-Discovery Experience with **Documill** Visual Search – Whitepaper 1 /4 2008-03-09  
White Paper **Documill** Visual Search – Enhance Your e-Discovery Experience ...  
Authors: Mika Könölä, Terho Laakso Date: 3/10/2008 Size: 448KB  
[http://www.documill.com/en/downloads/20080309\\_Documill\\_e-Discovery\\_wp.pdf](http://www.documill.com/en/downloads/20080309_Documill_e-Discovery_wp.pdf)  
View duplicates

**Presentation**  
Apache Tomcat) Whole SW solution is **Documill**'s own IP No 3rd party components, no  
blackbox approach – **Documill** has total control for the conversion ...  
Authors: Mika Könölä Date: 2/5/2008 Size: 1MB  
[http://www.documill.com/en/downloads/20080204\\_Documill\\_ContentProcessing.pdf](http://www.documill.com/en/downloads/20080204_Documill_ContentProcessing.pdf)

**Company Presentation**  
and affordable hardware platform(c) 2009 Documill, Inc. All rights reserved.Confidential  
material.8About **Documill**:**Documill** is an independent software vendor ...  
Authors: Mika Könölä Date: 11/23/2009 Size: 1MB  
[http://www.documill.com/en/downloads/20091119\\_Documill\\_VisualSearch.pdf](http://www.documill.com/en/downloads/20091119_Documill_VisualSearch.pdf)

**Visual Search Flyer**  
fax +358 9 2517 2202 sales@documill.com • www.documill.comwww.documill.com **Documill**  
Visual Search Fast ...  
Authors: Mika Könölä Date: 2/4/2008 Size: 1MB  
[http://www.documill.com/en/downloads/20080204\\_Documill\\_VisualSearch\\_2pager.pdf](http://www.documill.com/en/downloads/20080204_Documill_VisualSearch_2pager.pdf)

**Visual Search – Fast discovery of relevant information**  
search engines offer APIs thru which **Documill** Visual Search can leverage existing document  
indexes and search functionalities. **Documill** Visual Search – Whitepaper 4 /6 FAST ...  
Authors: Mika Könölä, Terho Laakso Date: 2/17/2008 Size: 1MB  
[http://www.documill.com/en/downloads/20080217\\_Documill\\_VisualSearch\\_wp.pdf](http://www.documill.com/en/downloads/20080217_Documill_VisualSearch_wp.pdf)

## Documill Visual Search enabled

Home > Search Results

All Sites

documill

Preferences Advanced

Result Type: 1-9 of 9 results

Any Result Type

Adobe PDF

Site: Any Site

www.documill.com/en

Author: Any Author

Mika Könölä

Mika Könölä, Ter...

Modified Date: Any Modified Date

Earlier

**Documill Dynamo**  
Interactive customer communication (c) 2009 Documill, Inc. All rights reserved.9About Documill:  
**Documill** is an independent software vendor (ISV ...  
Authors: Mika Könölä Date: 11/23/2009 Size: 1MB  
[http://www.documill.com/en/downloads/20091123\\_Documill\\_Dynamo.pdf](http://www.documill.com/en/downloads/20091123_Documill_Dynamo.pdf)

**Enhancing e-Discovery Experience with Documill Visual Search**  
Enhancing e-Discovery Experience with **Documill** Visual Search – Whitepaper 1 /4 2008-03-09  
White Paper **Documill** Visual Search – Enhance Your e-Discovery Experience ...  
Authors: Mika Könölä, Terho Laakso Date: 3/10/2008 Size: 448KB  
[http://www.documill.com/en/downloads/20080309\\_Documill\\_e-Discovery\\_wp.pdf](http://www.documill.com/en/downloads/20080309_Documill_e-Discovery_wp.pdf)  
View duplicates

**Visual Search Flyer**  
fax +358 9 2517 2202 sales@documill.com • www.documill.comwww.documill.com **Documill**  
Visual Search Fast ...  
Authors: Mika Könölä Date: 2/4/2008 Size: 1MB  
[http://www.documill.com/en/downloads/20080204\\_Documill\\_VisualSearch\\_2pager.pdf](http://www.documill.com/en/downloads/20080204_Documill_VisualSearch_2pager.pdf)

**Visual Search – Fast discovery of relevant information**  
search engines offer APIs thru which **Documill** Visual Search can leverage existing document  
indexes and search functionalities. **Documill** Visual Search – Whitepaper 4 /6 FAST ...  
Authors: Mika Könölä, Terho Laakso Date: 2/17/2008 Size: 1MB  
[http://www.documill.com/en/downloads/20080217\\_Documill\\_VisualSearch\\_wp.pdf](http://www.documill.com/en/downloads/20080217_Documill_VisualSearch_wp.pdf)

Filter 1 2 3 4 5 6 7 8 9

Filter 1 2 3 4

# User Experience Revolution

- Ensure that current search solution offers the best tools available for finding and accessing documents.
  - Focus on most valuable corporate assets; data hiding in enterprise documents
  - Significantly speed up document discovery experience
  - Enable integrated high-fidelity content viewing
- When core search functionality is reliable, you can make it easier for users to:
  - Verify search result relevancy
  - Improve insight into documents; page-level, document length
  - Find documents based on recognizable visual characteristics
  - Offer fastest access into content
  - Enable visually enriched, enjoyable user experience

# Visual Search User Experience

No need for content specific viewers  
View documents without having to download them

Thumbnail view provides immediate overview of the search results

Filter only pages with matching content visible

Simultaneous view of multiple documents instead of just one – faster relevancy verification

Instantly zoomable page previews from thumbnails

Keywords are highlighted even in the thumbnails

Ribbon view for understanding page context and quick preview of the whole document

The screenshot displays a search interface for 'dreamliner' using Microsoft Bing API 2.0. The search results are presented in a ribbon view, showing multiple document thumbnails simultaneously. The first result is titled '1. The Boeing 787 Dreamliner: More Than an Airplane' and includes a thumbnail with the text 'The Boeing 787 Dreamliner Work Than an Airplane'. The second result is '2. Dreamliner rolls out' with a thumbnail showing the aircraft on a runway. The third result is '3. Designing an airplane interior is important. The 787 Dreamliner ...' with a thumbnail showing the interior of the aircraft. The interface includes a search bar, a filter menu (pdf, ppt, doc), and a pagination control showing 56 results.

# Faster Way to Read Documents



Aviation and the Environment 

14 / 25

Navigation icons: back, forward, search, zoom in, zoom out, print, refresh

Search: dreamliner 

Found 3 results in 2 pages.

## The 747-8 is Cleaner, Quieter and More Efficient

The 747-8 delivers:

- 15%\* Reduction in fuel and CO<sub>2</sub>
- 28% Below 2008 industry limits for NOx
- 30%\* Smaller noise footprint



Advanced Materials

New Wing Design

Enhanced Flight Deck

787 Dreamliner Technology  
High-bypass Engines

Advanced Nacelles and Chevron Nozzles

\*Relative to the 747-400

COPYRIGHT © 2007 THE BOEING COMPANY

**LEADING**

Aviation and the Environment  
Our Commitment to the Future

James Yeh  
CEO, Boeing | Executive Vice President  
Global Operations | Chairman

1

Aviation and CO<sub>2</sub>

Global CO<sub>2</sub> | Global aviation CO<sub>2</sub>



2

CO<sub>2</sub>: A Global Challenge Needs a Global Solution



3

Aviation has made steady, significant progress



4

Our Strategy for Creating a Better Future

- Focus on a Clear Vision
- Technology, Innovation, Efficiency
- CO<sub>2</sub> and Fuel per Passenger
- Customer Efficiency
- A Global Network between manufacturers and airlines

# Mark & Manage Your Findings

## 11. Smiths Aerospace



Smiths Aerospace . Presentation to Investors . by: Dr John Ferrie. Group Managing Director. Cheltenham – Thursday 26 th May 2005

Navigation controls: Filter, 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20, Page 7 (click to zoom)



- Take actions from search results
- Mark pages, manage your findings stored in a server
- Collaborate with others
- Export findings as a new document

# Drag&Drop Bookmarking

Documill Visual Search Service for Internet

mika.konnola@documill.com  
[Edit account](#) | [Help](#) | [Sign out](#)

google web trends   [Prefs](#)

PDF  MS Powerpoint  MS Word

ALL RESULTS Results 1 - 8 of 1,570 hits for "google web trends"

### 1. New Trends on the Web

Theme . The one ginormous trend is "Power to the People" making the world more democratic. Information is flowing from common people to common people without the aid ...

Overview	What you should learn	Background	Theme	del.icio.us
<ul style="list-style-type: none"> <li>Social Bookmarking</li> <li>Tagging (a.k.a. Folksonomy) Flickr</li> <li>File sharing with BitTorrent et al.</li> <li>Geospatial Web - Google Maps, GoogleEarth</li> <li>Other Miscellaneous Items</li> </ul>	<ul style="list-style-type: none"> <li>What in the world is "Social Bookmarking"?</li> <li>What is tagging?</li> <li>What is Flickr?</li> <li>What does BitTorrent do for us?</li> <li>What possibilities arise from Google Maps?</li> </ul>	These new trends are being driven by these resources: <ul style="list-style-type: none"> <li>First, cheap internet access</li> <li>Cheap storage</li> <li>Open APIs (e.g. Weblogs (e.g., Google, Amazon))</li> <li>The cooperation of many, many people</li> </ul>	The one ginormous trend is "Power to the People" making the world more democratic. Information is flowing from common people to common people without the aid of established entities.	<ul style="list-style-type: none"> <li>What is del.icio.us?</li> <li>Del.icio.us</li> <li>What is your favorite web site?</li> <li>http://del.icio.us/feature2</li> <li>http://del.icio.us/feature1</li> <li>http://fresh.hennessy.com/del.icio.us.html</li> <li>Use del.icio.us via RSS</li> <li>Searching with http://del.icio.us/search/</li> <li>Find people with similar interests</li> </ul>

Filter 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

### 2. Trend Micro Web Gateway Security

Pre-Release Information. Aug 17, 2009 . Trend Micro Web Gateway Security . InterScan Web Security Virtual Appliance v5. Advanced Reporting and Management v1

Our Solution - Trend Micro Web Gateway Security	Software Virtual Appliance	What's New in IMS VA v5.0	Advanced Reporting and Management	Share Bookmark Organize	Systemic Dashboard - Real-time User Activity Monitoring
		<ul style="list-style-type: none"> <li>HTTPS and SSL scanning</li> <li>URL Filtering Enhancements</li> <li>Advanced Reporting and Management</li> <li>Systemic Dashboard</li> </ul>	<ul style="list-style-type: none"> <li>Stateful threat dynamic dashboard</li> <li>Over 50 pre-defined "Quick" Reports</li> <li>Dashboard viewer activity</li> <li>Real-time content monitoring and URL scan</li> <li>Social media monitoring</li> <li>Custom reports with Reports</li> <li>Policy management and synchronization for multiple sites</li> </ul>	<input type="button" value="Share"/> <input type="button" value="Bookmark"/> <input type="button" value="Organize"/>	

Filter 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Page 11 (click to zoom)

### 3. Web 2.0 trends & complexity

IBSG - 25 . Copyright © 2007 Cisco Systems, Inc. All rights reserved. Cisco Confidential . IBSG - 25 . Cisco Confidential . Common Twitter Application

Web 2.0 Defined	Paradigm and Complexity	Web 2.0 is a natural extension of a Services Oriented Architecture (SOA)	Characteristics of Social Networking	Online Youth Activities

Filter 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33

Web Statistics - 5 pages

Web Analytics addresses one of the most common business issues which companies face	Google Analytics: Quick Intro	Exporting	Historical Perspective: Rates of Adoption	The Long Tail in Perspective

- Choose relevant pages directly from search results and document listings
- Drag&drop pages to build collections
- Share collections with other users
- Export collections as new documents

# Documill Visual Search Suite

## 1. Search Result Set Previewer

- See what you have found without having to download them
- Overview of your findings prior drilling into those which look most promising
- Fastest way to find pages which matter most

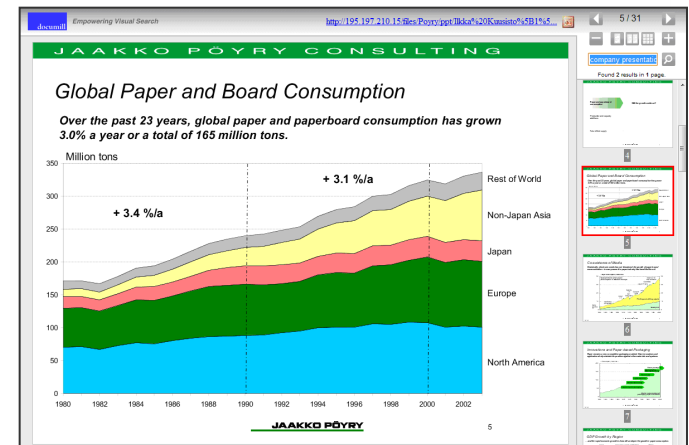
## 2. Document Viewer

- No need to download and open documents with 3rd party viewers; use browser instead
- Most efficient way into the content
- Secure way to view documents which might contain harmful payloads

## 3. Visual page-level bookmarking

- Mark pages you've found, and share your findings with others
- Create collections to organize your findings.
- Controlled sharing of collections enables collaborative search
- Create new documents from pages found

The screenshot shows a Google search interface with the query 'company presentation pöyry'. The search results are displayed in a preview format, showing a snippet of text from a document titled 'Jaakko Pöyry Group Company presentation 2006'. The preview includes a small thumbnail image of the document's cover page and a list of search results. The search results list includes 'Jaakko Pöyry Group company presentation', 'Pöyry presentation 2006', and 'Pöyry -let sales 1985-2005'. The search results are sorted by relevance and show the first 5 of about 5 results.



# Documill Explorer

Integrated document search, preview and page collections

The screenshot displays the Documill Explorer interface. At the top, there is a navigation bar with a search field containing "/Google Docs - joe.sample@leanfile.com/translate testing" and a search button. Below this is a menu bar with options like Organize, Views, Collections, Translate, and Viewer. The left sidebar shows a hierarchical file tree under "documill", including folders for "users", "groups", "tdc", and "Google Docs". The main area features a table of documents:

Name	Date modified	Size
100408_Space-Access10.pptx	7.4.2011 12:28	1 735 KB
361837main_10 - COTS Status Jun 17, 2009 r4.pptx	7.4.2011 12:27	5 524 KB
Corliss_CLARREO_Mission_Options_07092010.pptx	7.4.2011 12:27	9 855 KB
lostA2.pptx	5.4.2011 14:57	1 189 KB
The-Nature-of-Light-22.1.pptx	5.4.2011 15:08	916 KB

Below the table is a "Collections" section with "My Collections" and "Discovery" sub-sections. The right pane shows a preview of a presentation slide titled "SpaceX COTS System" with a description and features. At the bottom, there is a "Share" and "Export" section with a row of document thumbnails.

# Benefits

- Speed up the document discovery process → **Up to 7x faster discovery**
  - No need to download documents for screening
  - No need to carry out secondary searches within document viewers
  - Search and visually screen thru larger amount of data
- Save up to 3h/week → **Improved productivity**
  - According to IDC, Information Worker spends 9.5 h/week for searching information within enterprise
  - Most of the information is in documents spread around different document repositories or file shares
- Less downloading → **Significant bandwidth savings**
  - Visual Search provides immediate view of the content relevancy without forcing user to download the whole document
  - Avoiding unnecessary downloading can save up to 90% of bandwidth per document

# What's next?

- Overwhelming volume of data → People need to learn to co-operate to cope with it → **Enable collaborative search**
- Enterprises re-create document assets they already have, but just can not find → Users want to discover and re-use findings in their own creations → **Enable convenient re-use of existing assets**
- Search as a standalone solution does not work → Document repositories, asset management, workflows, intergated social networking → **Provide complete, integrated solutions**
- Less than 10% of companies use search for mining internal document repositories → Search is difficult to buy, nobody owns it. → **Offer Search as a Service (SaaS), easy to try & buy.**

# About Documill

- Documill is an independent software vendor (ISV) with solutions which...
  - enable browser-based access to MSOffice and PDF documents
  - empower large scale content processing solutions
- Core team has worked together over ten years in various mobile and digital media ventures.
- Core competency is server-side enterprise document processing.
- Headquartered in Espoo, Finland

