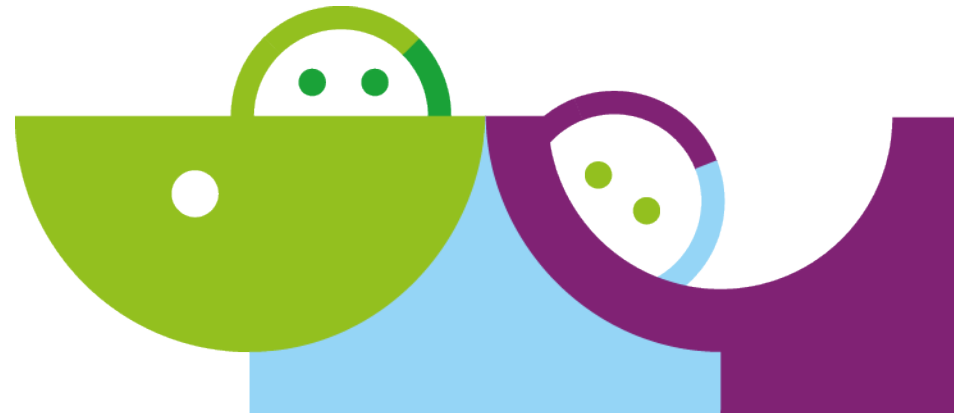


# Search Integration in WebSphere Portal - ein Überblick -

*Andreas Prokoph - XWebX Development - Lead Architect for Search, SME for Search and SEO*

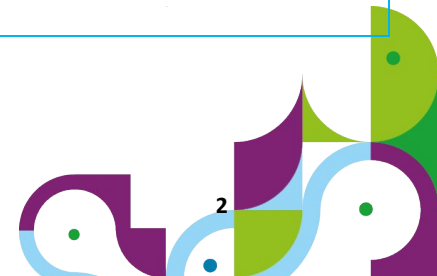


## Please Note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.



# Agenda

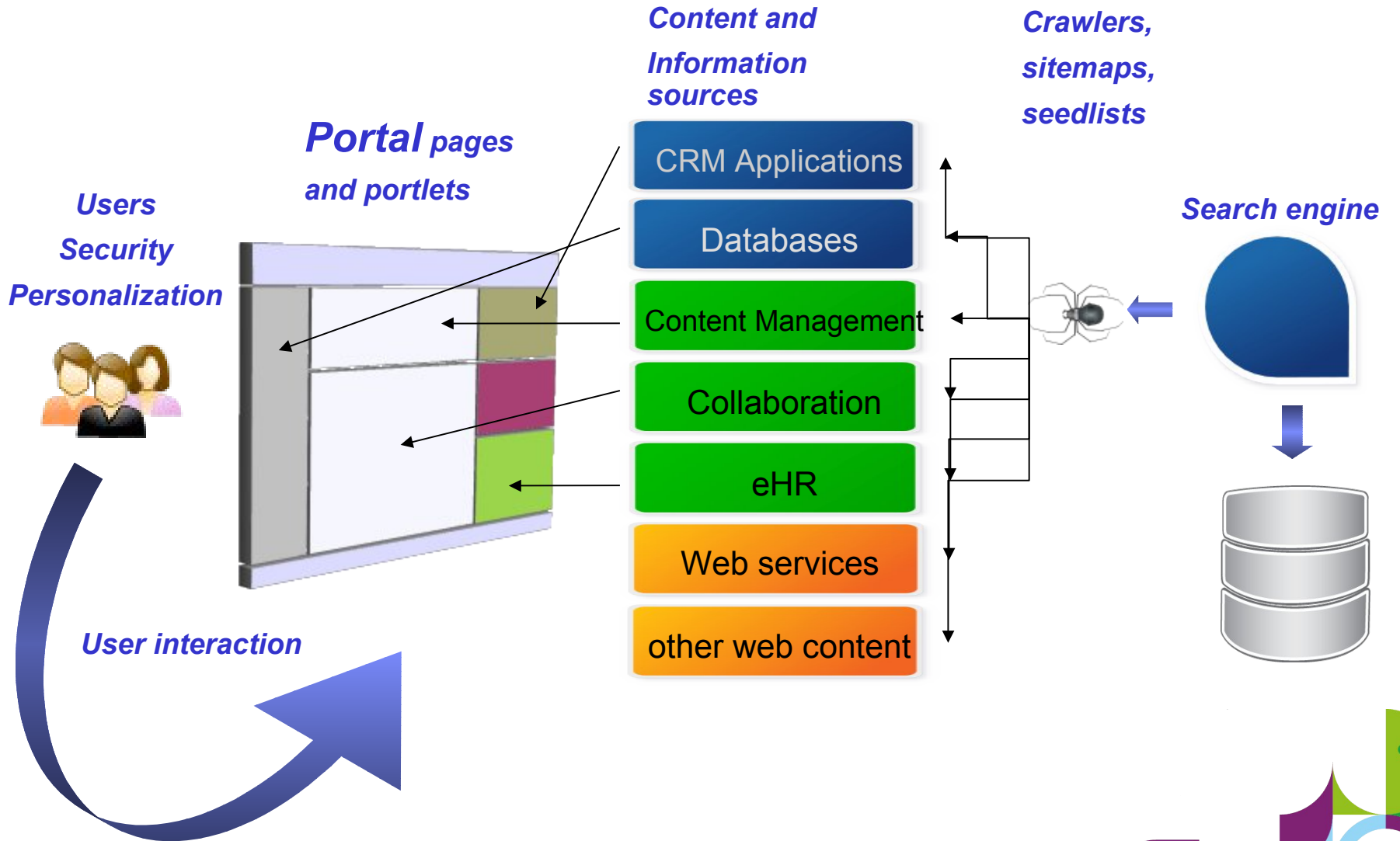
- Overview and general remarks
- Search integration frameworks in Portal
- Overview Portal Search
- Watson suite of Enterprise Search
- Summary



# Overview



# “Just crawl Portal and we're done” ... not exactly ...



## WebSphere Portal and search integration – in general ..

- **Q:** Can I integrate 'My' search engine with Portal?
- **A:** In theory: Yes. However some requirements must be met.
  - Portal provides frameworks to allow to integrate search engines:
    - IBM Seedlists – consume/crawl repositories efficiently
    - Security service – Puma REST APIs
    - Search UI
- Thus: '*In theory*' refers to whether or not that search engine in questions supports those frameworks
  - Consume IBM Seedlist to capture and process Portal and the like artifacts
  - Provide web interface to query the search service and return results as RSS or Atom feeds
- If your search engine of choice is Portal Search or IBM Watson Enterprise Search - then all requirements are met already



# Search integration frameworks



# What do we need?

- Access to Portal and web content repositories
- Security support
- Show items in the correct context of the Portal
- Search page and portlet(s)





## Access to Portal and web content repositories

- IBM Seedlists

- XML datastream
- Supports pagination
- incremental indexing – *list only changes since last visit*
- metadata and security information
- 'Actions' – add / update / delete
- URLs - distinguish between *fetch by crawler* and *display for end user*



# Security support

- Seedlist contains references to resources which the requesting user is entitled to
- “acls” entries contain unique IDs managed by Portal Access control
  - or: managed by IBM Connections (community memberships)
- Search engine may use *Puma REST service* calls to retrieve requesting user's entitlements and map against information stored in search index



# Show items in the correct context of the Portal

```
<atom:entry>
  <atom:id>15e0da37-0cdf-4dbb-a0a5-c367cf49cd7d</atom:id>
  <atom:link href="/wps/my poc/!ut/p/digest!W0HqDTtwoo7Plmv8c_ksNfWcmPa0BA5Temp....
  <atom:content src="/wps/wcm/myconnect/Template%20Page%20Content/Articles/List%20of
  <atom:title>List of Articles</atom:title>
  <atom:updated>2011-08-08T04:14:21+02:00</atom:updated>
  <wplc:action do="insert" />
  <wplc:acls>
    <wplc:acl>Z8eAe3BD6MO471BPAJMS609E6JMG6K1C8JM4CC1CCMMGC33PGJSK6L1E03OGCG1</wplc:acl>
    <wplc:acl>acl:allauthenticated</wplc:acl>
    <wplc:acl>Z9eAeL9CE3HDC2JD6MMG6JHDAMMG62RCCMM8CHPC2M061PJCMPU0P6SK623</wplc:acl>
  </wplc:acls>
  <wplc:field id="Name">List of Articles</wplc:field>
  <wplc:field id="secureContext">/mypoc</wplc:field>
  <wplc:field id="defaultContext">/poc</wplc:field>
  <wplc:field id="ContentPath">/Template Page Content/Articles/List of Articles</wplc:field>
  <wplc:field id="AuthoringTemplate">List</wplc:field>
  <wplc:field id="Modifier">uid=wpsadmin,o=defaultWIMFileBasedRealm</wplc:field>
  <wplc:field id="ExpirationDate">Jan 31 2010 14:00:00 CET</wplc:field>
  <wplc:field id="EffectiveDate">Feb 09 2011 05:34:02 CET</wplc:field>
  <atom:published>2011-02-09T05:27:03+01:00</atom:published>
  <atom:summary>This list enumerates all articles created in context of this page.</atom:summary>
</atom:entry>
```

**“display URL”  
for the end user  
(handled by content handler)**

**“crawler URL”  
let crawler focus solely on  
the content**



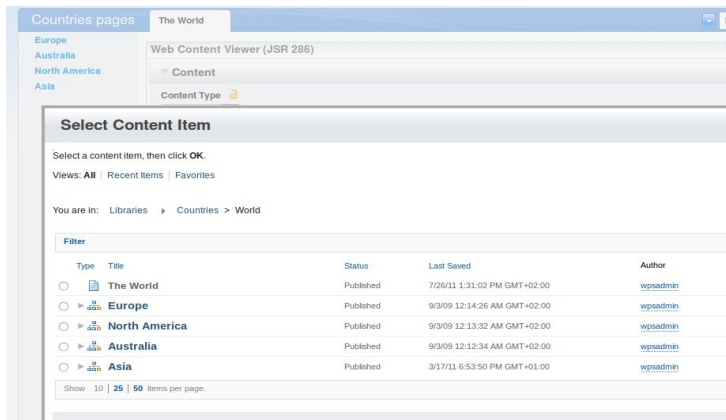
# Example of handling content managed by IBM WCM

*Wiring of WCM content with the Portal site infrastructure ....*

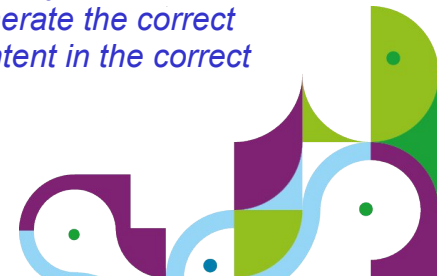
2. .. and is associated with a 'sitearea' in a WCM library



1. Portal page is of type 'content' ...



*.. what is achieved now, is that when a content object needs to be rendered in the Portal, a 'content handler' will then check what sitearea the content belongs to, and then retrieve the appropriate Portal page ID it is associated with. This allows to then generate the correct URL so display the content in the correct context of the Portal.*



# Overview - Portal Search



## Built-in Portal Search gives you ...

- **IBM Search Features**

- Delivers [federated search](#) across Portal content sources and any linked sites
- Search honors portal [security](#)
- [Relevance](#) ranking
- [Suggested Links](#)
- Integration of [Tagging and Rating](#)
- Custom [search scopes](#) for filtering
- [Document formats](#) (more than 250)
- [Summaries](#) and [Meta Data](#)

- **Administration**

- Crawl & Index: Portal and WCM content, IBM Connections, standard web sites
- Search Seedlist support optimizes source indexing
- Programmatic access for Portal search through REST Services
- Pre-configured search collections and crawlers OOTB (Portal, WCM, WCM Authoring)



## Search Center ootb

All Sources

More than **10** results found in **All Sources** for **portal**

Showing 1 - 10 Previous Page | Next Page

| Relevance | Title   | Person | Date            |
|-----------|---|--------|-----------------|
| ★★★★★     | <b>Manage Virtual Portals - Virtual Portal Manager</b><br><b>Summary:</b> Administration portlet to manage virtual portals<br><a href="#">Edit keywords for Suggested Links</a>     |        | 1/23/13 8:54 PM |
| ★★★★★     | <b>Virtual Portals - Information Portlet</b><br><b>Summary:</b> Display an information message contained in a JSR286 .JSP file<br><a href="#">Edit keywords for Suggested Links</a> |        | 1/23/13 8:29 PM |
| ★★★★★     | <b>Portal Analysis - Information Portlet</b><br><b>Summary:</b> Display an information message contained in a JSR286 .JSP file<br><a href="#">Edit keywords for Suggested Links</a> |        | 1/23/13 8:29 PM |
| ★★★★★     | <b>About IBM WebSphere Portal - About WebSphere Portal</b><br><b>Summary:</b> Displays Version and Copyright Statement<br><a href="#">Edit keywords for Suggested Links</a>         |        | 1/23/13 8:29 PM |
| ★★★★○     | <b>inlineedit.js</b><br><a href="#">Edit keywords for Suggested Links</a>   |        | 1/24/13 1:44 AM |
| ★★★★○     | <b>MoveBlogToBeDeleted</b><br><a href="#">Edit keywords for Suggested Links</a>   |        | 1/24/13 1:45 AM |

### IBM developerWorks : Lotus : Articles

#### developerWorks

Deliver an exceptional mobile web experience using WebSphere Portal and IBM Worklight, Part 3: Implementing automatic single sign-on with Worklight and WebSphere Portal

More and more, enterprises are providing multi-channel support to their web channel communities. This article explains how IBM WebSphere Portal and IBM Worklight enable enterprises to create Worklight applications capable of automatically logging a user in upon startup, while also using single sign-on (SSO) to log the user into a WebSphere Portal server on the same host at the same time. This makes it easy for mobile apps to display portal pages customized to the user from within the application.

Deliver an exceptional mobile web experience using WebSphere Portal and IBM Worklight, Part 4: Using IBM Web Experience Factory with Worklight to create hybrid applications

IBM Web Experience Factory is a powerful model-based tool for rapidly developing custom portlets and web applications that are optimized for mobile and desktop devices. IBM Worklight is a complete mobile enterprise application platform for delivering native, hybrid, and web applications. This article shows how you can use



## Hints 'n' tips

- *Crawling a portal site and about security filtering*
- *Crawl Portal and WCM – filtering available*
- *Customize Search Center and search page*
- *Integrating with IBM Connections*
- *Search quality tuning options*
- *Faceted search*
- *Multi-language search*
- *Date range constraints*















# Crawling a portal site – which crawler(s) should I use?

- OOTB we have a Portal search collection configured with two content sources:
  - **Portal Site** - pages and portlets
  - **WCM library(-ies)**

Search service: **Default Search Service** Search collection: **Default Search Collection**

**Content Sources**

Use the icons and buttons below to create a content source, refresh the content source information, view schedulers, start and stop crawling, verify the URL address of a content source, edit a content source, and delete a content source. **Scopes that reference this collection and its content sources can return different results.**

|                       |           |          |                          |                          |        | Page 1 of 1   |
|-----------------------|-----------|----------|--------------------------|--------------------------|--------|---|
| Content Source Name   | Documents | Run Time | Last Run                 | Next Run                 | Status |   |
| WCM Content Source    | 25        | 00:00:33 | May 16, 2013, 12:50 CEST | May 17, 2013, 03:00 CEST | Idle   |      |
| Portal Content Source | 116       | 00:00:02 | Apr 26, 2013, 18:29 CEST | May 17, 2013, 01:00 CEST | Idle   |      |

Page 1 of 1

- For **public pages** it is OK to use the **web crawler** feature
- If the majority of information is served through WCM, use the **WCM Site crawler only**
  - efficient incremental crawling supported – *allows for short crawling intervals*
  - Quality of search and search result increases
    - *focus is on the content and its own meta data*
    - *Deleted items removed instantly when signaled with next scheduled crawl*



## Seedlist crawler – what about filters?

- Traditional web crawler filters often not applicable
  - *since relying on URL patterns is often useless*
- Filtering of seedlist entries themselves is not meaningful
  - *if less entries delivered than were requested, crawler could assume 'last page' encountered*
- **Solution:** let the respective application decide what should and what should not be indexed
  - *Mark **portlets** with a 'searchable' flag as appropriate (default is 'true')*
  - *Same goes for **WCM content** items*



## Apply filters on portlets

- Identify the portlet(s), select 'Configure' and add or change the property "INCLUDE\_IN\_SEARCH\_INDEX" to 'true' or 'false' as required

**Manage Portlets**

Search by: Title contains Search:

**Portlets** Click Copy to create a duplicate of the portlet. Click Configure to set titles, descriptions and parameters. Click Delete to remove the portlet from your portal. Click Assign Access to allow others to work with the portlet.

| Title                        | API Type | Unique name                          | Provided | Remote portlet | Status |                  |
|------------------------------|----------|--------------------------------------|----------|----------------|--------|------------------|
| Web Content Libraries        | IBM API  | ibm.portal.Web.Content.Admin.Lib     |          |                |        |                  |
| Web Content Authoring        | IBM API  | ibm.portal.Web.Content.Authoring     |          |                |        |                  |
| Web Content Subscribers      | IBM API  | ibm.portal.Web.Content.Admin.Sub     |          |                |        |                  |
| Web Content Viewer (JSR 286) | JSR 286  | ibm.portal.Web.Content.Viewer.Jsr286 |          |                |        | <b>Configure</b> |
| Web Content Syndicators      | IBM API  | ibm.portal.Web.Content.Admin.Syn     |          |                |        |                  |

Page 1 of 1



**Manage Portlets**

**Configure portlet: Web Content Viewer (JSR 286)**  
Web module: ilwwcm-localrenderingportlet-jsr.war

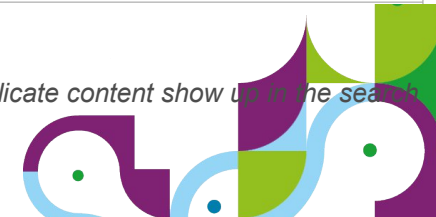
**Preference and Values** Enter a new preference and value pair in the blank fields to create a new preference for this portlet, or click Delete to remove a preference and value. Click OK to keep your changes or Cancel to quit.

New Preference:  New value:

| Preference              | Value               |  |
|-------------------------|---------------------|--|
| INCLUDE_IN_SEARCH_INDEX | false               |  |
| SECTION_STATES          | true,true,true,true |  |
| WCM_BROADCASTS_TO       | WCM_LINKING_DYNAMIC |  |

Page 1 of 1

**Note:** the 'Content Viewer' portlet is disabled by default to thus avoiding having duplicate content show up in the search result list.



## Filtering of WCM content – control now given to the business users

- *Single Authoring template with the appropriate flag – need to be selectively put on or taken off*
  - *In the sample Authoring template the default is 'selected'.*

Option 1

The screenshot shows the IBM WebSphere Portal authoring interface for an 'Untitled' document. The document was created by 'wpsadmin'. The interface includes a toolbar with buttons for 'Save and Close', 'Add To Project...', 'Hide Hidden Fields', 'Add Workflow', 'Manage Elements', and 'Close'. Below the toolbar are tabs for 'Content' and 'Properties'. The 'Properties' tab is active, showing the following details:

- Type:** Content
- \*Location:** Select Location (Web Content / Articles (First child))
- Authors:** Add Authors (wpsadmin)
- Owners:** Add Owners (wpsadmin)

---

**Content Properties**

- Content Template:** Web Content / Article
- Search Collection Visibility:**  Include this item in search collections. Changes here will take effect next time the collection is updated.

---

**Profile**



## Filtering of WCM content – control now given to the business users

- *Less 'hidden' and more obvious to the editors: two flavors of the Authoring template, distinguished by a meaningful name or title that suggests if that content is either searchable or not.*

Option 2

|         |                        |                           |                     |                 |          |         |
|---------|------------------------|---------------------------|---------------------|-----------------|----------|---------|
|         | Content                | Collaboration             | Messaging           | Personalization | wpsadmin | Actions |
| Welcome | Web Content Management | WCM Library Import/Export | Web Content Preview |                 |          |         |

### New Content

Select an authoring template, then click **OK**.

Views: **All** | Recent Items | Favorites

Libraries ▶ Web Content

Filter

| Type                             | Title   | Status    | Last Saved                   | Author |
|----------------------------------|---|-----------|------------------------------|--------|
| <input type="radio"/>            | <b>Article</b>                                      | Published | 5/15/12 6:48:55 AM GMT+02:00 |        |
| <input checked="" type="radio"/> | <b>Article (Article for non-searchable content)</b> | Published | Today 4:10:29 PM GMT+02:00   |        |



## Filtering of WCM content – *file attachments*

1) `SearchService.SearchSeed.ExcludeFileAttachments=[false|true]`

- 'true' prevents file resource component attachments from being included in search
- 'false' (default) enables to search for file resource elements

2) `SearchService.SearchSeed.excludeExtensions`

- Define a list of file name extensions that you want to exclude from search results

example: `SearchService.SearchSeed.excludeExtensions=avi,mpeg,zip`

See: [http://www-01.ibm.com/support/knowledgecenter/SSYJ99\\_8.5.0/wcm/wcm\\_config\\_search.dita](http://www-01.ibm.com/support/knowledgecenter/SSYJ99_8.5.0/wcm/wcm_config_search.dita)



## .. or maybe even this?

The screenshot shows the IBM WebSphere Portal interface. At the top, there is a navigation bar with 'Welcome', 'MySearch page', and 'Site Map'. On the right, there are links for 'wpsadmin', 'Actions', 'Log Out', and the IBM logo. Below the navigation bar, there is a section titled 'You might be interested in' with a dropdown arrow. Underneath, there is a link for 'Search Administration - Information Portlet' with 'Keywords: search' and an 'Edit' link. Below this is a search bar with 'All Sources' selected and a search input field containing 'search'. A message states 'More than 10 results found in All Sources for search'. The main content area lists search results with summaries, dates, and suggested links. On the right side, there is a 'Tags' section with a list of tags and a 'View as' option set to 'cloud'. Below the tags is an 'Experts' section with three profiles: Hahn, Roberta J.; Prokoph, Andreas; and Meyer, Brian C. Three blue star-shaped callouts are overlaid on the page: one pointing to the search bar area with the text 'Changed formatting (results.xml)', one pointing to the tags section with 'Add Tag Cloud and Tag Results portlets', and one pointing to the experts section with 'Add External Results portlet to show Additional (related) information'.

Welcome MySearch page Site Map wpsadmin Actions Log Out ? IBM

You might be interested in

[Search Administration - Information Portlet](#)  
Keywords: search  
[Edit](#)

All Sources search

More than 10 results found in All Sources for search

**IBM Intranet | Enterprise search results**  
Summary: Skip to main content. Skip to search filters. Skip to search box.  
Date: 2012-08-09T11:20:50.866Z  
[Edit keywords for Suggested Links](#)

**LegalFlow Software for Text Searching. Legal Document Management.**  
Summary: LegalFlow Software for Text Searching. Legal Document Management. Take Care of Business: Selected e-Forms allow for both manual and automatic addition and auto-population of data, such as legal information, from multiple and simultaneous sources.  
Date: 2012-09-27T07:54:00.062Z  
[Edit keywords for Suggested Links](#)

**MySearch page - Search Center**  
Summary: This portlet allows performing a scoped search and viewing Portal search results  
Tags: Search\_center,  
Date: 2012-10-08T11:23:05.000Z  
[Edit keywords for Suggested Links](#)

**w3 Directory | Organization | Corporate Headquarters**  
Summary: Skip to main content. The access keys for this page are:. Additional accessibility information for w3.ibm.com can be found on the w3 Accessibility Statement page.  
Date: 2012-08-09T11:20:08.239Z

Tags

All tags

App\_overview Collaboration\_ST  
Collaboration\_welcome Content\_authoring  
Content\_preview Content\_welcome  
Getting\_started Homepage Library\_explorer  
Messaging Poersonal;ization\_business\_rules  
Search\_center Sitemap Tag\_center  
Unified\_task\_list UTL WCM  
View as **cloud** | list

Experts

**Hahn, Roberta J.**  
IBM Certified Executive Project Manager - GWPE -  
Search PDLT  
IBM CHQ, BT/IT CIO  
CHICAGO  
IL  
United States

**Prokoph, Andreas**  
WebSphere Portal - Search Architect for Portal and WCM, SME for Search  
IBM Software Group, IBM Collaboration Solutions  
BOEBLINGEN  
Germany

**Meyer, Brian C.**  
Executive Architect - Global Workforce & Web Process Enablement (GWPE)  
IBM CHQ, BT/IT CIO  
BOULDER  
CO  
United States

Changed formatting (results.xml)

Add Tag Cloud and Tag Results portlets

Add External Results portlet to show Additional (related) information

## Integration with IBM Connections search

- **Option 1** - use IBM Connections seedlists to crawl Connections repositories
  - One seedlist per resource type, e.g. Wikis, Blogs, Profiles, Files, etc..
- **Options 2** - integrate IBM Connections Search with Portal's Search Center
  - use 'Remote Content Search Service' (RCSS) feature to register Connections search as a search service
  - IBM Connections and WebSphere Portal leverage IBM Search REST APIs
  - Note: for security filtering it is required to set search service configuration parameter 'ExternalSecurityResolverUrl' to something like: `https://host:port/wikis/seedlist/authverify/getACLTokens`

See: [http://infolib.lotus.com/resources/connections/4.0/doc/en\\_us/ic4\\_p4.html#connectors\\_portlets\\_search\\_rcss](http://infolib.lotus.com/resources/connections/4.0/doc/en_us/ic4_p4.html#connectors_portlets_search_rcss)





## You need to utilize custom meta data for search?

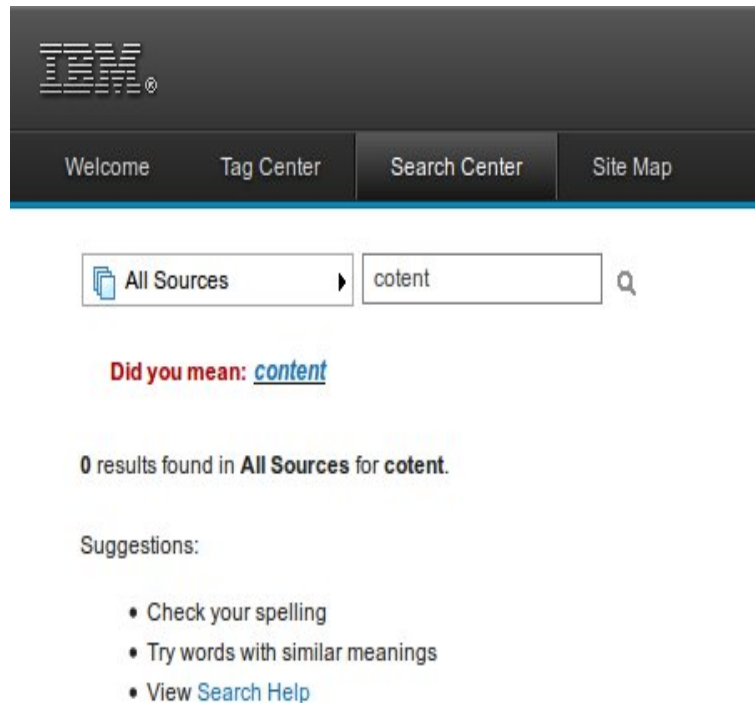
- Publish custom fields through WCM seedlist
  - Add the following custom properties to the WP ConfigService resource environment using the WebSphere Integrated Solutions Console
    - `wcm.config.seedlist.metakeys=metakey1,metakey2`
- Display in search result
  - either by developing custom portlet
  - Use DDC to render search result
  - or modify results.xsl to capture and output your custom meta data
- Note: metadata is published as “*META\_*metakey1”
  - e.g. instead of field name “Region” the field will be available as “META\_Region”

[http://infolib.lotus.com/resources/portal/8.0.0/doc/en\\_us/PT800ACD004/wcm/wcm\\_dev\\_search\\_enableseed.html](http://infolib.lotus.com/resources/portal/8.0.0/doc/en_us/PT800ACD004/wcm/wcm_dev_search_enableseed.html)



# Search quality improvements

- User query correction – aka 'Did you mean?'



The screenshot displays the IBM WebSphere Portal search interface. At the top, the IBM logo is visible. Below it, a navigation bar contains links for 'Welcome', 'Tag Center', 'Search Center', and 'Site Map'. The search bar shows 'All Sources' as the selected source and 'cotent' as the search term. A magnifying glass icon is to the right of the search bar. Below the search bar, a red message reads 'Did you mean: [content](#)'. Underneath, it states '0 results found in All Sources for cotent.' and 'Suggestions:' followed by a bulleted list:

- Check your spelling
- Try words with similar meanings
- View [Search Help](#)



## Search quality – IBM LanguageWare

- Dictionary based language support
- Supported languages:
  - Language Identifier
    - Afrikaans, Arabic, Balinese, Basque (Euskera), Bulgarian, Catalan, Chinese (Simplified and Traditional), Czech, Danish, Dutch, English, Finnish, French, German, Greek, Hebrew, Hungarian, Icelandic, Irish (Gaelic), Italian, Japanese, Korean, Malay, Norwegian Bokmal, Norwegian Nynorsk, Polish, Portuguese, Romanian, Russian, Spanish, Swedish, Tagalog, Thai, Turkish, Vietnamese

For details see:

[https://publib.boulder.ibm.com/infocenter/ieduasst/imv1r0/index.jsp?topic=/com.ibm.iea.lrw/lrw/7.2/Overview/1-LanguageWare\\_Technology/player.html](https://publib.boulder.ibm.com/infocenter/ieduasst/imv1r0/index.jsp?topic=/com.ibm.iea.lrw/lrw/7.2/Overview/1-LanguageWare_Technology/player.html)



# Search quality improvements .. and the story goes

There are three options available to **im**prove search quality and thus overall search experience:

- Use the **Suggested Links** portlet and capability of Portal Search  
*Available since WebSphere Portal V6.1*
- Change the **default search behavior** to require all keywords to be present in every returned document, rather than being sufficient if at least one keyword is present – move from 'OR' to 'AND'  
*Available since WebSphere Portal V7*
- Apply a **boost factor** for the document in cases when the keywords appear in certain 'fields' (metadata) within the document  
*Available since WebSphere Portal V8.0.0.1 CF09*

See also: [http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Improving\\_search\\_quality\\_in\\_Portal\\_Search](http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Improving_search_quality_in_Portal_Search)



## Details – field boosting – search service configuration

Name: boostingSettings

Value: 

```
{"phraseBoost": {"Enabled": "true"},  
"fieldBoost": [{"field": "title", "boost": 3.0} ,  
                {"field": "description", "boost": 3.0},  
{"field": "keywords", "boost": 2.0}]}
```

*fieldboost*: *sample provided for default/common found metadata fields*

Can also include any other metadata fields (with string based values)  
“boost” should be specified in a range between 1.0 to 10.0, and should  
be used with care (suggested to stay in the range between 1.0 and 3.0).

*phraseboost*: not mandatory, could add value – though very much language dependent

*Note: Does not require re-indexing!*



## Portal Search – Multi-language support

CF05

- Today:
  - *The Search Center filters by language automatically based on user's preferred language*
- Requirement: allow to disable language filtering in the Search Center portlet
  - *today: manually update a script file – with the potential to be updated with the next CF installed (documented in a technical article)*
- Language information however is still required for query text analysis
  - *either use 'preferred language' of the user*
  - *or the search collection's default language*

See: [http://www-01.ibm.com/support/knowledgecenter/SSHRKX\\_8.5.0/mp/admin-system/config\\_search\\_multi.html](http://www-01.ibm.com/support/knowledgecenter/SSHRKX_8.5.0/mp/admin-system/config_search_multi.html)



# Portal Search – multi-language support

- *The following configuration options apply to the Search Center portlet configuration settings*
- *The Portal admin can switch on/off language filtering*  
`SEARCH_LANG = [PREFERRED | ALL]`
  - PREFERRED - the preferred language of the user is used to filter the search results. This is the default and matches today's behavior
  - ALL - There is no language specific result filtering applied
- The Portal admin can further determine what language information is used *for analyzing the search query text*:

`QUERY_LANG = (PREFERRED | COLLECTION)`

- PREFERRED - the preferred language is used
- COLLECTION – the language of the collection is used
- The **default value** depends on the value of `SEARCH_LANG`
  - `QUERY_LANG=PREFERRED`, if `SEARCH_LANG=PREFERRED`
  - `QUERY_LANG=COLLECTION`, if `SEARCH_LANG=ALL`





## Portal Search - Date and date range constraints

- You can search documents by **date** and **date ranges**. The used date formats can be changed by the Portal administrator. Dates can be specified by using the format yyyy-MM-dd. You can also specify the time as well as the date by using the format yyyy-MM-dd, hh:mm. The following example shows which search terms are supported:
  - **Date:** `update_date:"yyyy-MM-dd"` or `update_date:"yyyy-MM-dd, hh:mm"`
    - Finds all documents that were updated on the specified day, or
    - Finds all documents that were updated on the specified day and time
  - **Date range:** `#update_date::>=2014-10-14<2014-10-19`
    - Finds all documents that were updated between the two specified dates: documents that were updated between and including October 14, 2014 and October 19, 2014 (excluded) would be returned. Range queries support the following operators:
      - > (greater)
      - >= (greater equal)
      - < (less than)
      - <= (less or equal)
- Example: “find all status reports from 2014”
  - `query:"status reports" #creation_date::>=2014-01-01<2015-01-01`

See: [http://www-01.ibm.com/support/knowledgecenter/SSHRKX\\_8.5.0/mp/admin-system/srrcfgsrv.c.dita](http://www-01.ibm.com/support/knowledgecenter/SSHRKX_8.5.0/mp/admin-system/srrcfgsrv.c.dita)





## Portal Search administration task to enable date constraints



- The following parameters on the search service are used:
  - ***dateFieldPattern***  
specifies a regular exp that is used to check if a field name is a date field  
default: `.*date$`
  - ***dateFormat***  
A date format specified using the ICU SimpleDateFormat pattern  
default: `yyyy-MM-dd`
  - ***dateTimeFormat***  
A date format with a time part specified using the ICU SimpleDateFormat pattern.  
default: `yyyy-MM-dd, HH:mm`
  - ***dateFormatLocale***  
The locale that should be used for the date parsing given as a ICU ULocale LocaleID.  
default: Portal default locale for local search. For remote search the default Locale of the JVM
- See: [http://www-01.ibm.com/support/knowledgecenter/SSHRKX\\_8.5.0/mp/admin-system/srrcfgsrvc dita](http://www-01.ibm.com/support/knowledgecenter/SSHRKX_8.5.0/mp/admin-system/srrcfgsrvc dita)



# Portal Search - facets are available

Filtering widgets



**Tags**

- aix 21
- unix 20
- aix\_and\_unix 12
- tools\_and\_utilities 8
- application\_development 8
- system\_administration 8
- speaking\_unix 6
- java\_technology 4
- shell 4
- linux 4

**Popularity**

22085 100000

**Rating**

1 5

**Categories**

- Java technology 54

## IBM Explorative Portlet

Search

At least 22085 views X Rating 1 and above X

Applied filters removable controls

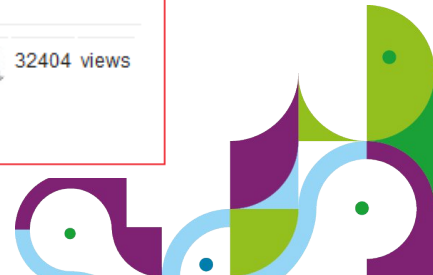


Main results view table



Showing 1 - 10 of 171 Previous | Next

| Title  | Last Update       | Rating | Popularity  |
|--|-------------------|--------|-------------|
| <b>Recommended reading list: DB2 for Linux, UNIX, and Windows database administration</b><br>Samuel Poon   Tags: administration, all, and, are, available, books, database, db2, db2luw1, for, general, here, linux, list%253A, listdb2, reading, recommended, unix, windows<br><Strong>Summary:</Strong> Learn about IBM's DB2; for Linux, UNIX, and Windows; with this reading list, compiled especially for the database administrator community. This popular article is updated to include the latest content that has been published for DB2 9 and 9.5. There is also an <a href="http://www.ibm.com/developerworks/db2/library/techarticle/dm-0509poon2/">http://www.ibm.com/developerworks/db2/library/techarticle/dm-0509poon2/</a> ; application development reading list; that is available as part of this series. | November 20, 2008 | ★★★★★  | 29528 views |
| <b>Enterprise search with PHP and Apache Solr</b><br>Martin Streicher<br><Strong>Summary:</Strong> Discover how to combine an enterprise-worthy search engine &#8212; Apache Software Foundation's Solr &#8212; with your PHP application.   | January 15, 2008  | ★★★★☆  | 74429 views |
| <b>Crossing borders: REST on Rails</b><br>Bruce Tate<br><Strong>Summary:</Strong> Earlier articles in the <a href="#">Crossing borders</a> series introduced Ruby on Rails as an explosively popular   | August 1, 2006    | ★★★★☆  | 32404 views |



## How to enable Portal Search to exploit Search Facets

- What we have available today:
  - Updated Search REST documentation now including the search facets
    - [http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Search\\_REST\\_API\\_specification\\_javadoc8](http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Search_REST_API_specification_javadoc8)
  - Sample code, portlet and data - for demonstrating use of search facets
    - <https://greenhouse.lotus.com/plugins/plugincatalog.nsf/assetDetails.xsp?action=editDocument&documentId=0DA010D7A05CABD785257B0E00324E20>
    - Note: you need a Greenhouse ID to get the sources
  - Finally: the Portal Wiki article “that ties the knots together”
    - [http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Leverage\\_Search\\_Expierence\\_based\\_on\\_Portal\\_Search\\_REST\\_API](http://www-10.lotus.com/ldd/portalwiki.nsf/dx/Leverage_Search_Expierence_based_on_Portal_Search_REST_API)



these are the 'basics'  
.... need more?



# Watson Search

(Explorer Editions V10,  
Content Analytics V3.5)



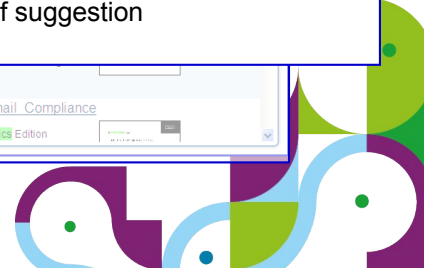
# Adding IBM Content Analytics (IBM Enterprise Search) to the picture

- **Enhanced End-User Experience**
  - High-performance faceted navigation
  - Saved searches
  - Search profiles
  - Document previews
  - And more...
  
- **Enhanced Administrative Experience**
  - Scalability improvements
  - Incremental indexing support
  - Reduced resource requirements
  - More flexible scale-out & HA
  - New relevancy tuning options
  - And more...

The screenshot shows the IBM WebSphere Portal search interface. The search query is "ibm search analytics". The results page displays a list of documents with columns for Source, Date, Title, and Thumbnail. A Facet Tree on the left shows filters for "date" with options for 2009 (9), 4 (9), 20 (9), and 15 (9). A Dynamic Facet Chart is also visible. A type-ahead search dropdown is shown, suggesting terms like "state", "state street", "stated", "state department", "states", "statement", "stater", "statesboro", "states", "stateed", "stater", "statesboro".

| Term             | Estimated results |
|------------------|-------------------|
| state            | 5                 |
| state street     | 440               |
| stated           | 2                 |
| state department | 610               |
| states           | 1                 |
| statement        | 90                |
| stater           | 4                 |
| statesboro       | 2                 |
| states           | 2                 |
| stateed          | 2                 |
| stater           | 2                 |
| statesboro       | 2                 |

**Type ahead search:**  
 Suggests queries based on index content and past queries  
 Shows estimated results count as part of suggestion  
 Customizable by Search Administrators



# Search Center Integration of WCA

Search Center
Welcome Search Center Site Map

All Sources
Q

Close Search

Title:

Author:

**3 results found in All Sources for +author:Eitan.**

|               | Relevance Title   | Person        | Date             |
|---------------|---|---------------|------------------|
| Showing 1 - 3 | <div style="display: flex; align-items: flex-start;"> <div style="margin-right: 5px;"> <span style="font-size: 0.8em;">● ● ● ● ●</span> <span style="font-size: 0.8em; color: #4a86e8;">🌐</span> </div> <div> <a href="#">Integrating IBM Lotus Sametime with the IBM Lotus Quickr Search REST service</a><br/>                     Summary: Skip to main content. In our example, the open source Apache Abdera client is used to parse the Atom feeds in the SearchFeedParser class. The summary element can be used as hover text, and the image element is the icon for the scope.                 </div> </div>  | Eitan Shapiro | 1/22/08 2:00 AM  |
|               | <div style="display: flex; align-items: flex-start;"> <div style="margin-right: 5px;"> <span style="font-size: 0.8em;">● ● ● ● ●</span> <span style="font-size: 0.8em; color: #4a86e8;">🌐</span> </div> <div> <a href="#">IBM WebSphere Developer Technical Journal: Introducing the Search and Indexing API in WebSphere Portal V6.0</a><br/>                     Summary: Skip to main content. In the sections that follow, we explore the different packages, explain the underlying concepts, and detail code snippets from the sample search portlet included in this article. In the same way, if you know the name of the collection you want to query, you can call genericSearchService.getSearchable(your collection name).                 </div> </div>                          | Eitan Shapiro | 10/25/06 2:00 AM |
|               | <div style="display: flex; align-items: flex-start;"> <div style="margin-right: 5px;"> <span style="font-size: 0.8em;">● ● ● ● ●</span> <span style="font-size: 0.8em; color: #4a86e8;">🌐</span> </div> <div> <a href="#">Unleashing the power of WebSphere Portal V6 Search with the Portal Search Toolbox</a><br/>                     Summary: Skip to main content. You can use the Scope Search portlet to provide a query search user interface with three elements: the text area for query input, the search button to execute the query, and a search scope selection list. For example, you can choose to display ten results in the Display Results portlet, but, in the background, use the ResultSet Size to request at least one hundred results.                 </div> </div> | Eitan Shapiro | 7/30/07 3:00 AM  |

Number of results per page: 10 [20](#) [30](#) [50](#)

Showing 1 - 3



## IBM Content Analytics V3 Search portlets (incl. sample code)

The screenshot displays the IBM WebSphere Portal search interface. At the top, there is a navigation bar with links for Home, Administration, Applications, Search Center, Enterprise Search, and Tag Center. A search bar on the right contains the text 'wpsadmin' and 'Actions', and a search input field with the text 'Search'. Below the navigation bar, there are tabs for 'Getting Started', 'Features', 'TestPage1', 'TestPage2', and 'TestPage3'. The main content area is divided into two portlets. The left portlet, titled 'ica\_search', contains a search input field with the text 'IBM' and a search button. Below it is a 'FacetTreePortlet' showing a tree view of search results. The tree is organized by date, with categories for 2009 (12), 2001 (2), 8 (1), 11 (1), 2002 (1), 2005 (1), and 2008 (1). There is also a 'directory' category. At the bottom of the facet tree, there are radio buttons for 'New search' and 'Add to search', and 'Search' and 'Clear' buttons. The right portlet, titled 'ResultListPortlet', displays search results for 'IBM'. It shows 'Results 1-10 of 17' and a 'Show Details' button. The results are listed in a table with columns for date, title, and snippet. The first result is 'OmniFind Tech Overview' dated April 20, 2009. The second result is 'ecmoverview.pdf' dated April 20, 2009. The third result is 'content-analyzer-datasheet.pdf' dated April 20, 2009. The snippet for the third result reads: 'Enterprise content management solutions Better decisions, faster IBM Content Analyzer Actively leverage ... In fact, it's estimated that anywhere from 50 to 80 percent of potentially relevant business information resides in ... intelligence. IBM Content Analyzer software can help you convert content into business insight in support of an information ... in your organization Helps reduce customer attrition and improve customer satisfaction and loyalty with better ... opinions Mitigates recall and warranty costs by proactively identifying quality control issues Helps detect ... earlier and more accurately Harness scattered content, operate more intelligently Content Analyzer software ... tion Management Architecture (UDMA) framework, IBM Content Analyzer 8.4.2 uses sophisticated natural language ... supporting integration ...'.



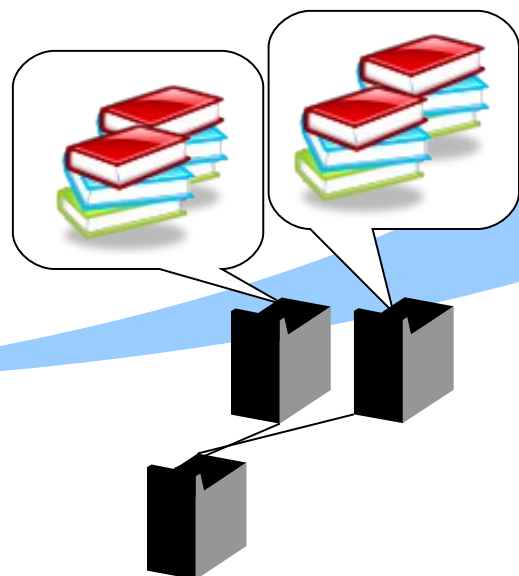
# Seamless Scale-out options

Content Analytics with Enterprise Search offers 3 types of system configuration based to the volume of data

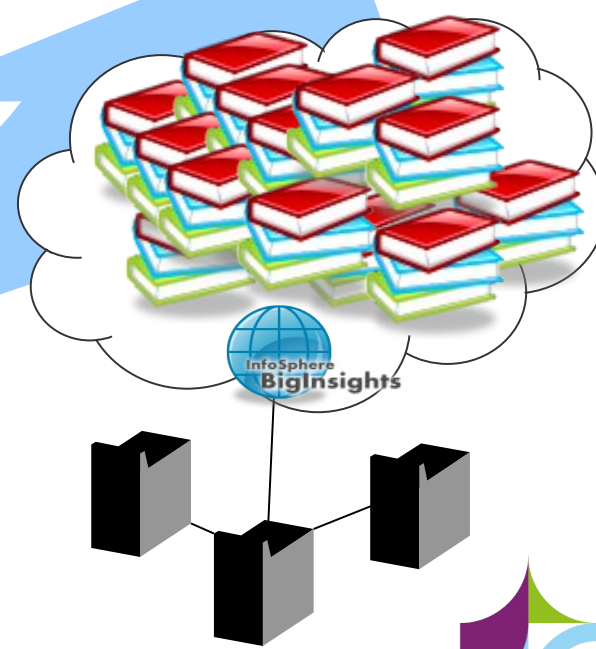
Analysis of small datasets can be done on a single machine



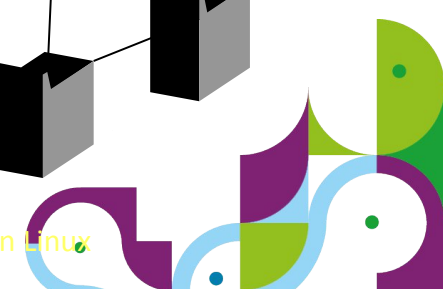
Production system can be deployed to 1 to N servers



Production system analyzing big data can utilize BigInsights



\* BigInsights is supported only on Linux



# IBM Watson Content Analytics Enterprise-Wide Content Reach

## Collaboration

- IBM Case Manager V5.0
- IBM Lotus Connections 2.5.0.1 and 2.5.0.2
- IBM Lotus Domino DM 6.5.1, 7.0,
- IBM Lotus Domino R7, R8 and R8.5
- IBM Lotus Quickr (NSF & J2EE), 8.1, 8.2 and 8.5
- Lotus Web Content Management 6.1, 6.1.5
- IBM WebSphere Portal 6.1 and 6.1.5

## Content Management

- IBM Case Manager V5.0
- IBM Content Manager Enterprise Edition 8.4, 8.4.1, 8.4.2
- FileNet Content Services 5.4, 5.5
- FileNet P8 Content Manager 4.0, 4.5, 4.5.1 and 5.0
- Hummingbird DM 5.1.0.5 with SR6 and 6.0.4
- EMC/Documentum 6.0 and 6.5
- CA-Datacom 10
- Open Text Livelink Enterprise Server 9.6, 9.7 and 9.7.1

## Data Management

- DB2 for iSeries 5.4 and 6.1
- DB2 UDB for Linux, UNIX, Windows 9.1, 9.5 and 9.7
- DB2 for z/OS 8.1, 9.1
- IMS 10.0 and 11.0.1
- Informix Dynamic Server 11.10 and 11.50
- Microsoft SQL Server 2005, 2008
- MySQL 5.0
- Oracle 10g and 11g
- Software AG Adabas
- Sybase 11.9.2, 12.0

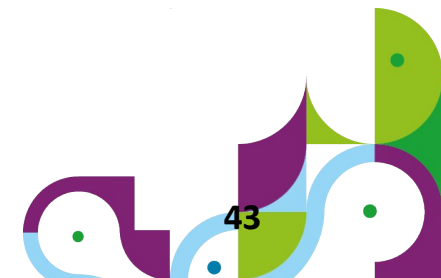
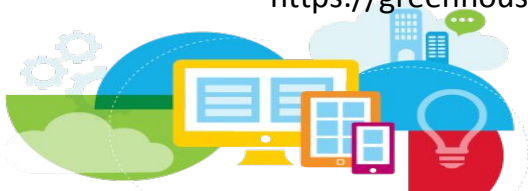
## Miscellaneous

- Microsoft Exchange Server 2003
- Microsoft Windows SharePoint Services 3.0
- SharePoint Server 2003 SP2, 2007, 2010 Farm Support
- Windows file systems: 2003 and 2008 servers
- Network News Protocol Newsgroup
- UNIX file systems
- VSAM for z/OS 1.4
- Web (HTTP or HTTPS)



## References

- **Portal Search – Quick Starter Guide**  
[http://www-10.lotus.com/ldd/portalwiki.nsf/dx/apoSearchapo\\_Quick\\_Starter\\_Guide\\_for\\_WebSphere\\_Portal](http://www-10.lotus.com/ldd/portalwiki.nsf/dx/apoSearchapo_Quick_Starter_Guide_for_WebSphere_Portal)
- **Watson Content Analytics / Explorer**  
<http://www-01.ibm.com/support/docview.wss?uid=swg27037837>  
<http://www.ibm.com/smarterplanet/us/en/ibmwatson/explorer.html>
- **IBM Digital Experience Solutions**  
<http://www-01.ibm.com/software/collaboration/digitalexperience>
- **WebSphere Portal and IBM Web Content Manager Information Center Wiki**  
<http://www-10.lotus.com/ldd/portalwiki.nsf/>
- **IBM Digital Experience Demonstrations:**  
<http://www.youtube.com/user/IBMXWebX>
- **IBM Digital Experience Developer**  
<http://developer.ibm.com/digexp>
- **IBM Software Business Solutions Catalog**  
<https://greenhouse.lotus.com/catalog/>



# Questions?

