



CERN Search & Social for the Enterprise Web experience

Andreas Wagner - CERN IT/OIS
Bruno Silva De Sousa - CERN IT/OIS

HEPiX Spring Meeting
Oxford, UK
Mar 23, 2015



Outline:

- Status of CERN Web Services
 - Overview & Status
 - Web Site Life Cycle Management, Web Analytics
- CERN's Enterprise Social Networking System
 - Motivation & purpose
 - Feature overview: profiles, social networking, microblogging, suggestion systems and discussion forums
- CERN Search
 - Status
 - Enterprise search
 - New features

WEB

Status of CERN Web Services

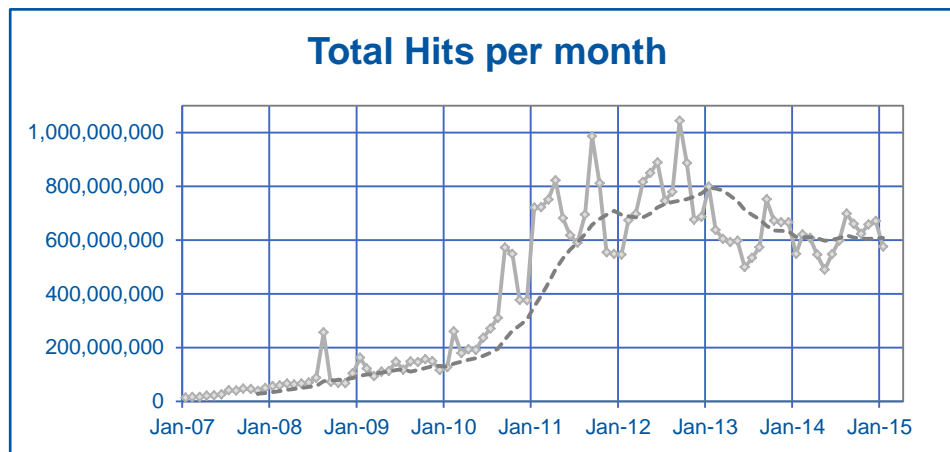
- Mandate: “Support for authoring and publishing content on the web.”
- About 200 Webservers (many cloud based)
 - Windows IIS with DFS Storage,
 - Apache with AFS,
 - Drupal,
 - Sharepoint,
 - TWiki (IT/PES)
 - Java MiddleWare On Demand site (IT/DB)
 - APEX workspaces (IT/DB)
- Wide range of technologies for dynamic content: Perl, PHP, Python, Django, .NET framework, ...

CERN Web Services in Numbers

- **Number of sites:**

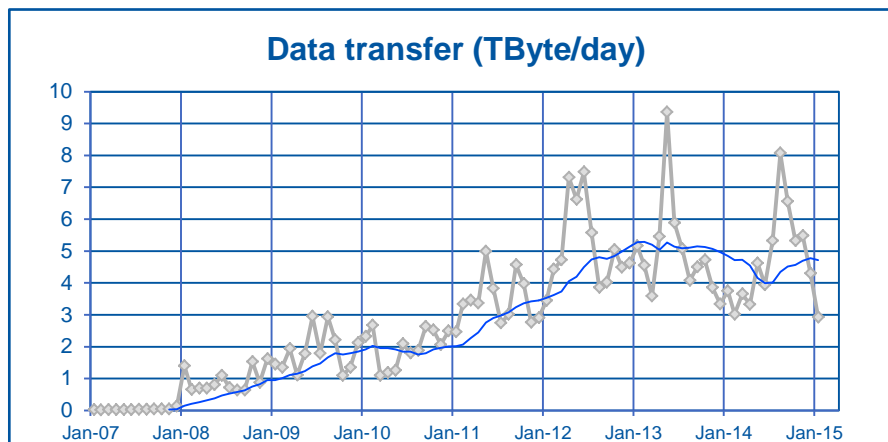
Total 14455 sites on ~200 Web servers (2015-03-22)

Number of sites	
Centrally hosted (IIS & DFS)	6100
AfsGateway (Apache & AFS)	3742
Sharepoint	2790
Drupal	836
Twiki	520
J2EEApp	279
APEX	186
Other	2
Total	14455



- **Web traffic:**

- Sep 2014: 700 million hits
- Sep 2014: **8 TByte / day**

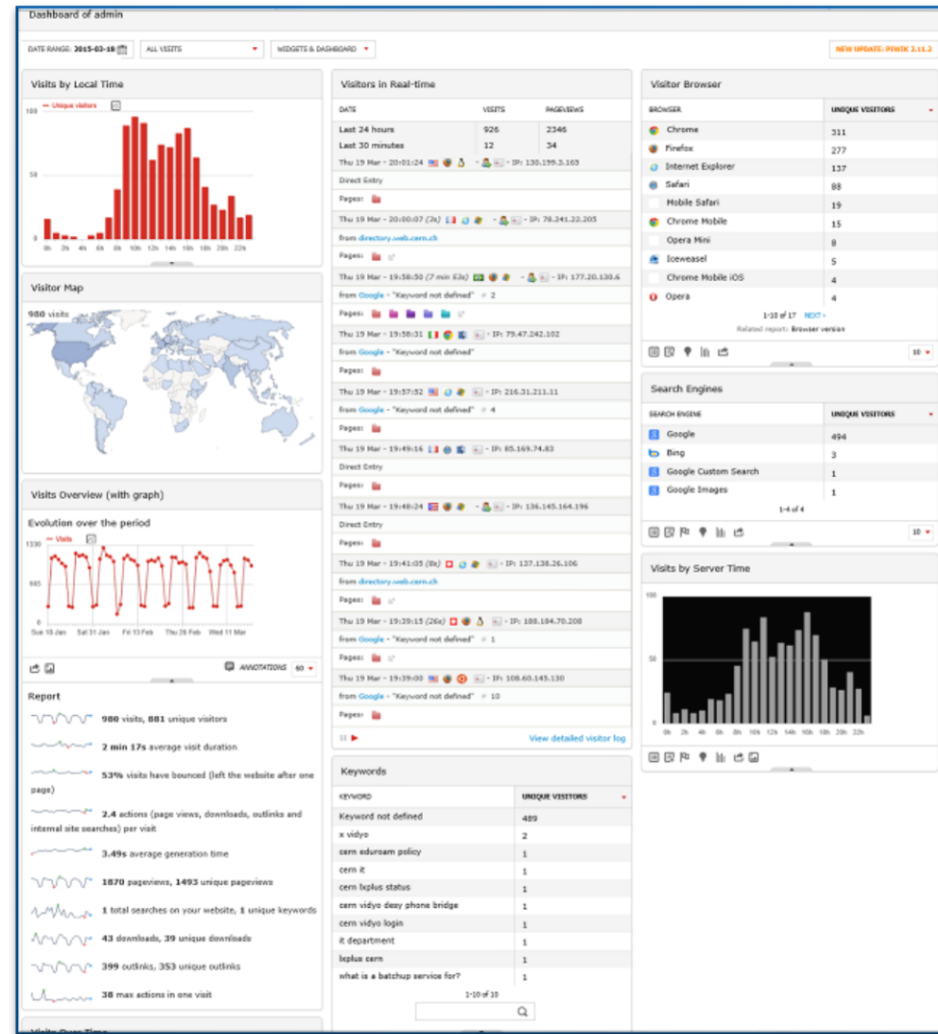


New features (I)

- **Web Site Life Cycle Management**
 - Website life-cycle as for computer accounts
 - Personnel sites will be blocked/deleted when owner leaves
 - Official Websites will be reassigned to supervisor
 - Will help to prevent orphaned sites
 - Will allow to clean-up unused web-space
 - spring-cleaning:
 - ~300 websites with 0 visit within last 3 years
 - ~3000 websites with <10 visit within last 3 years

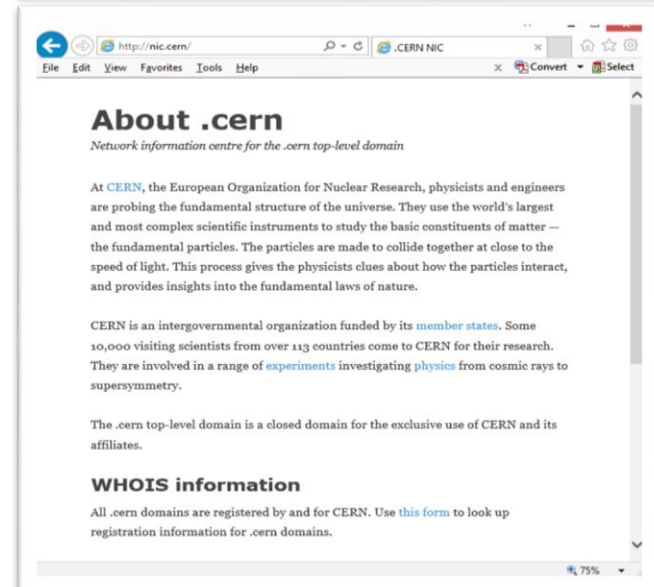
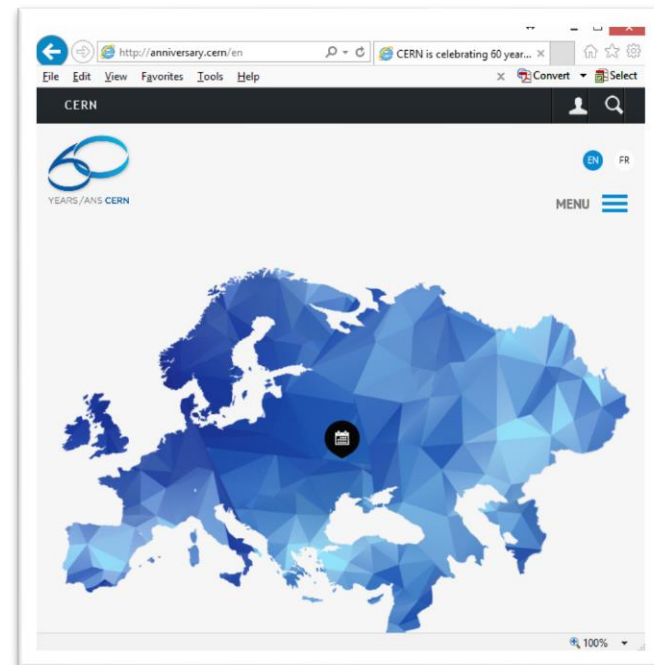
New features (II)

- Web Analytics (Piwik)
- Popular open-source Web analytics package
- Accessible for site owners



New features (III)

- Support for new top level domain **.cern**
 - <http://anniversary.cern>
 - <http://nic.cern>
- Used by CERN communications team for major CERN public events and outreach activities



CERN Web Services in 2015

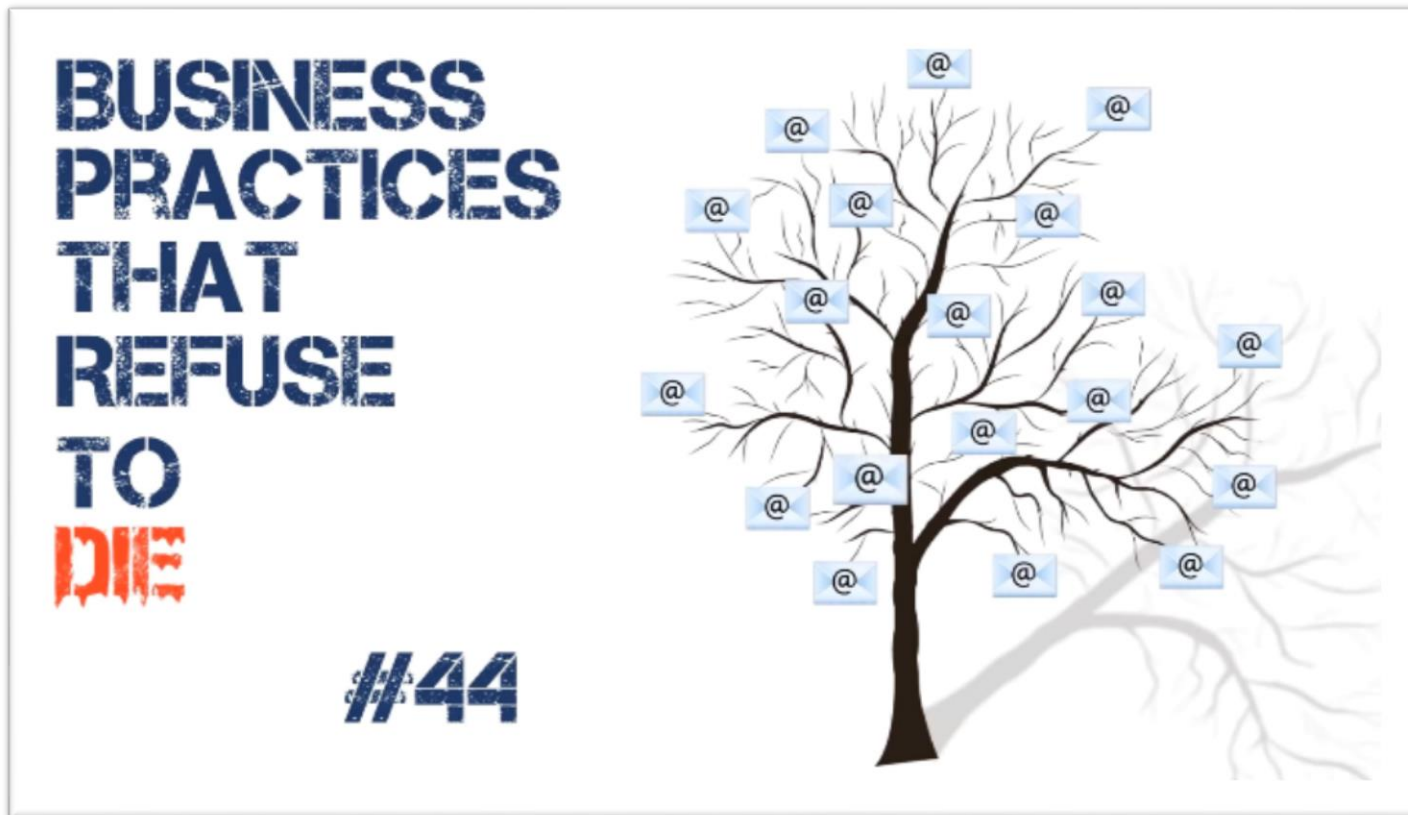
- Migration Sharepoint: 2010 → 2013
- Migration WebAFS: SLC5 → SLC6
- Migration IIS 7.5 → 8.5
- Drupal 8 pilot
- WebAFS CC7

Outline:

- Status of CERN Web Services
 - Overview
 - Web Site Life Cycle Management, Web Analytics
- **CERN's Enterprise Social Networking System**
 - Motivation & purpose
 - Feature overview: profiles, newsfeed, social communities
- CERN Search
 - Background, Enterprise Search
 - Document Collections Status
 - New features

SOCIAL

CERN's ESN: Motivation (I)



- **Business Practices That Refuse To Die #44: Email Trees**
https://www.youtube.com/watch?v=QlqA_YKeboc

Enterprise Social Networking (I)

- We receive too many emails
- Push & Pull channels
- Reply all syndrome
- Knowledge locked in mailboxes
- To avoid communication & knowledge silos
- Shift of general communication habits

Enterprise Social Networking (II)

- Use of technologies to create value by connecting members of an organization
- Similar to public social media but for organizations (LinkedIn, Twitter, G+, Facebook)
- Unlock opportunities to communicate, share knowledge and collaborate

Enterprise Social Networking (III)

- ESN components
 - Social Profiles and Newsfeeds
 - Communities and Workgroups
 - Blogs and Wikis
 - Suggestion System and Social Voting (“Likes”)



Enterprise Social Networking (IV)

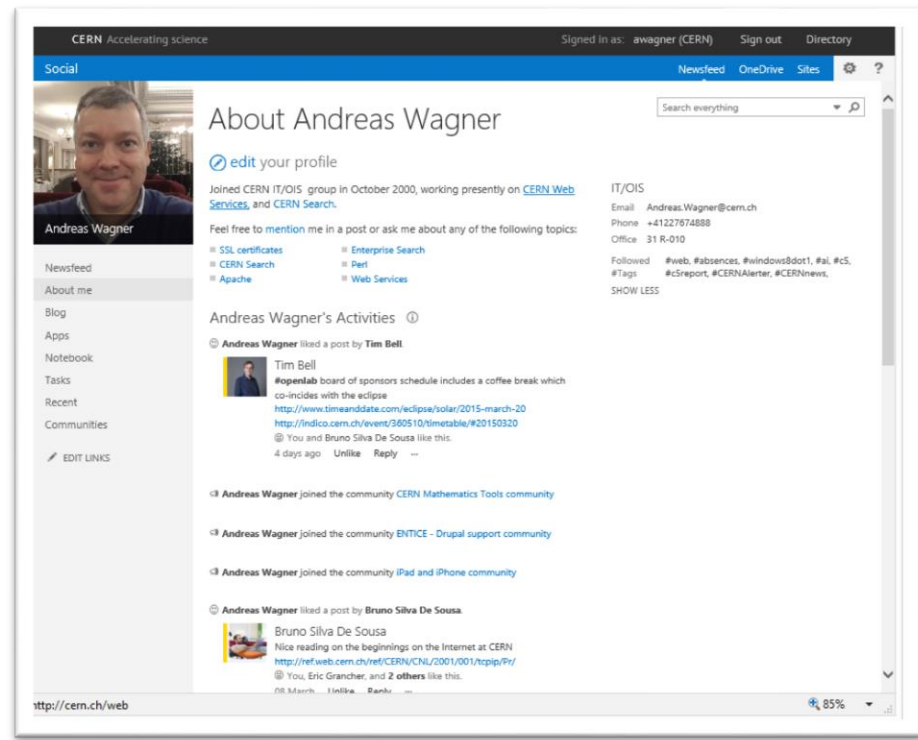
- Enterprise Social Networking
 - 80% cultural change of communication and interaction
 - 20% technology
- ESN offers a lightweight communication channel
 - a single place to stay up-to-date about what going on within organization
 - Pull channel i.e. people follow what they are interested in
 - Facilitate information sharing (across hierarchical boundaries)
- Conversations taking place in ESN allow
 - To collaborate more efficiently and allow others to contribute
 - To create content that becomes searchable and accessible to many

Social @ CERN

- CERN's ESN implementation
 - SharePoint 2013
 - support for all standard compliant of browsers
- Social for CERN
 - Social Newsfeed & Profiles
 - Social Communities
 - Integrated with OneDrive, Office Web Apps, Search
 - Programmatic access

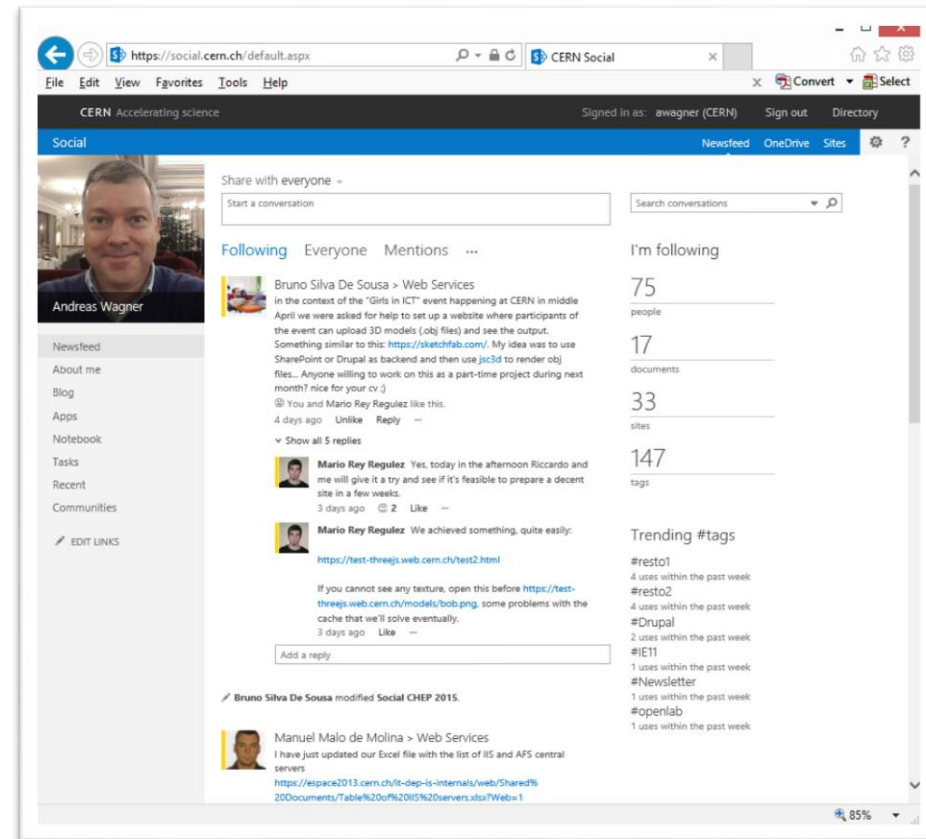
Social Profiles

- On first access profile is created (for anyone with CERN account)
 - Add information about yourself, picture, areas of expertise
 - Configure notifications and shared activities



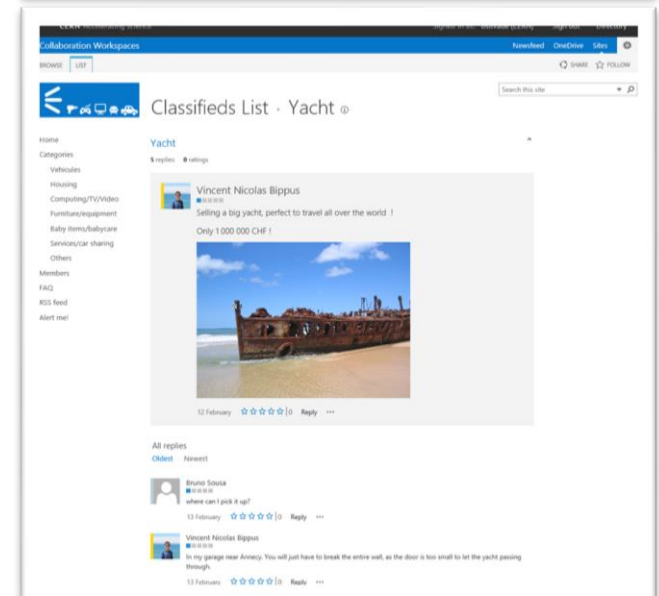
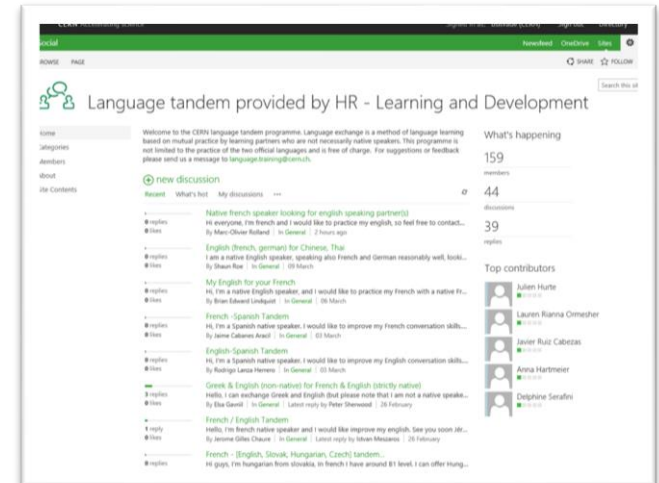
Social Newsfeed

- What you're **Following**
 - people, #tags, documents or workspaces
 - By default you follow your team members
- What **Everyone** is saying
- Your **Activities**
- Your **Mentions**
- Your **Likes**



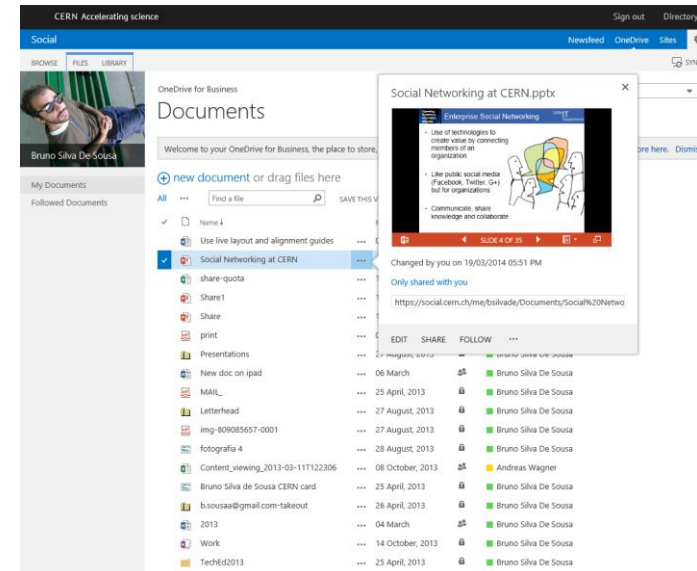
Social Communities

- Forum experience
 - Topic, Q&A oriented
- Can be opened also for external accounts
- e.g. IT Lightning Talks, Software Development, Language Tandems, CERN Market, ...



CERN Social: integrated with ...

- OneDrive (for business)
 - Place to easily share and collaborate on files
 - Can be 'Shared with Everyone' (Anyone logged on CERN SSO including Federation)
- Office Web Apps
 - Browser-based for Word, PowerPoint, Excel, OneNote and PDF files
 - Allows concurrent editing



Social Search

- Find content
 - Context tabs
 - In-page previews
 - Documents
 - Conversations
 - Visual Refiners
- Features coming to next version of CERN Search

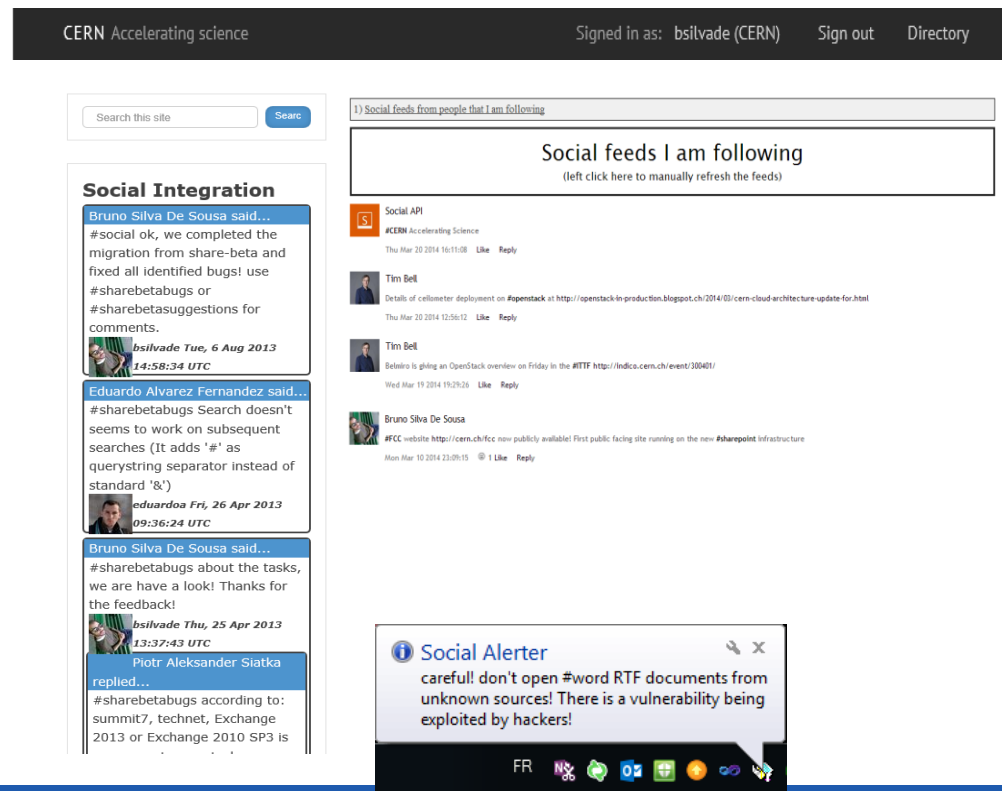
The screenshot shows a search interface for 'esn' on the CERN Social platform. The search results are displayed in a list format with various filters and options. The left sidebar includes filters for 'Result type' (Newsfeed post, OneNote, PDF, PowerPoint, Word), 'Author' (Bruno Silva De Sousa, System Account, Piotr Aleksander Siatka, CT), and 'Modified date' (One Year Ago, Today). The main content area shows several search results, including a document titled 'sparglight-wp-1202-altimeter-esn' and a video titled 'The Rise of Enterprise Social'. A right-hand sidebar features a preview for 'Social Networking at CERN ITTF' by Bruno Silva de Sousa, dated 28th March 2014. The interface includes navigation tabs for 'Everything', 'People', 'Conversations', and 'Videos', and a 'Preference for results in English' dropdown.

The screenshot shows a search interface for 'social' on the CERN Social platform. The search results are displayed in a list format with various filters and options. The left sidebar includes filters for 'Tags' (#social, #teched2013, #sharepoint2013, #somehash, #pow), 'Author' (Bruno Silva De Sousa, Bruno Sousa, Tim Bell, Andreas Wagner), and 'Created' (One Year Ago, Today). The main content area shows several search results, including a document titled 'social social-beta regis...' and a video titled 'adding posts to Social progr...'. A right-hand sidebar features a preview for 'Social Community Site' with 6 members and popular discussions. The interface includes navigation tabs for 'Everything', 'People', 'Conversations', and 'Videos', and a 'Preference for results in English' dropdown.



CERN Social - programmatic access

- To allow integration with existing Web applications.
- REST API to interact with Social
 - SSO (with CORS), NTLM, Kerberos, Basic authentication
- JavaScript API developed to ease integration
- Drupal modules available (read/write)
- Plugin for CMF
- Plans to integrate Social with common CERN Web Toolbar



The screenshot displays the CERN Social web interface. At the top, it says "CERN Accelerating science" and "Signed in as: bsilvade (CERN)". Below this is a search bar with the text "Search this site" and a "Search" button. The main content area is divided into two columns. The left column, titled "Social Integration", contains a list of recent posts from users like Bruno Silva De Sousa and Eduardo Alvarez Fernandez. The right column, titled "Social feeds I am following", shows a list of posts from users like Tim Bell and Bruno Silva De Sousa. At the bottom right, there is a "Social Alerter" notification box with the text "carefull! don't open #word RTF documents from unknown sources! There is a vulnerability being exploited by hackers!".

Outline:

- Status of CERN Web Services
 - Overview & Status
 - Web Site Life Cycle Management, Web Analytics
- CERN's Enterprise Social Networking System
 - Motivation & purpose
 - Feature overview: microblogging, profiles, social networking, suggestion systems and discussion forums
- **CERN Search**
 - Background, Enterprise Search
 - Document Collections Status
 - New features

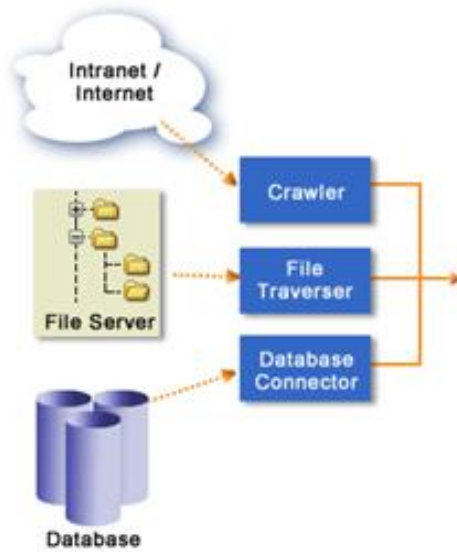
SEARCH

CERN Search

- A central Search solution to provide
 - **for users**
 - Single entry point for searching information on wide range of content sources at the same time
 - **for service providers**
 - Search backend service
 - Indico, TWiki, Drupal, SharePoint, JACOW, E-group archives, (EDMS)
- **History**
 - Start of project in 2006:
based on commercial product FAST ESP
 - Now: FastSearch for SharePoint 2010
 - In preparation Search SharePoint 2013

Enterprise Search

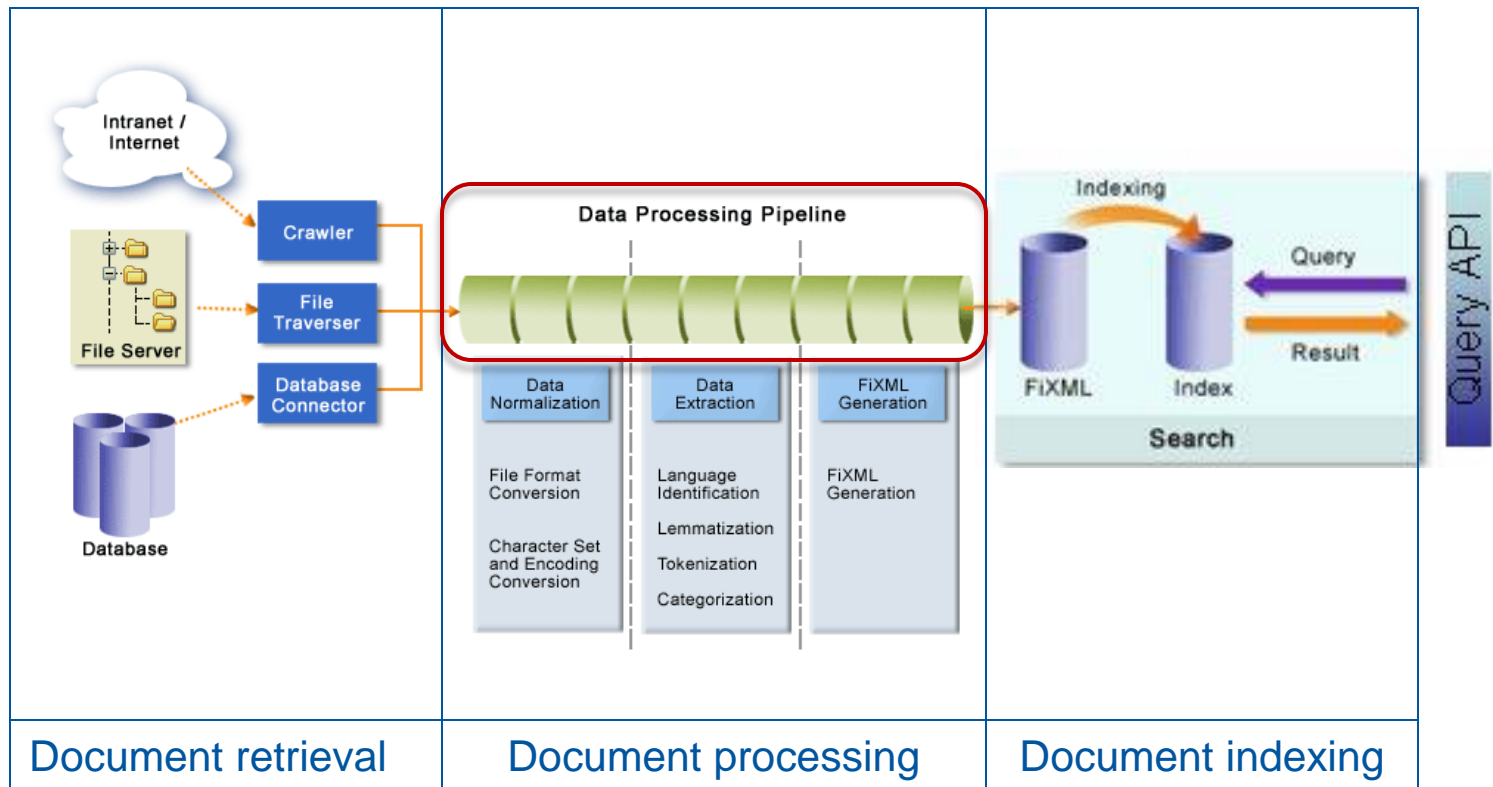
- Wide range of document sources:



- Web Pages (Public & Protected)
- File systems
- Databases
- Directories (People and Places)
- Document repositories (CDS, EDMS, Indico, ...)
- Structured Content Management System Data
 - SharePoint, Drupal, TWiki

- Variety of meta data
- Different Access Protection Schemes
- Different retrieval methods and frequencies

Document Content Flow



Indexing Protected Content I

- 🔒 To allow indexing protected content we need to
 - Retrieve the document
 - Search engine needs access to document,
 - Crawlers are not necessarily SSO aware
 - Obtaining document ACLs
 - To be able to decide who is allowed to find a document
 - Often not trivial since most systems answer the question:

“Has a given user the right to access a given document?”

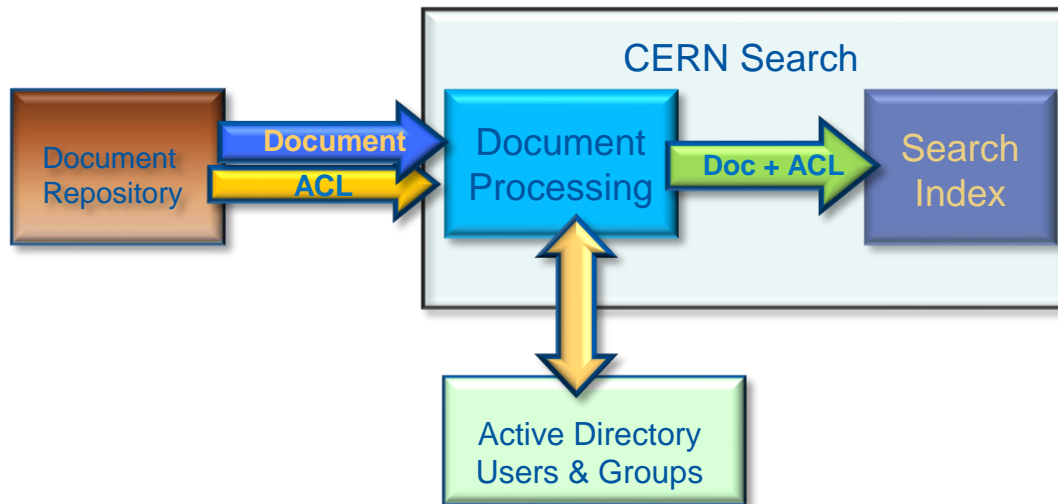
and not

“Tell me who has access to a given document?”






This is due to often complex permission models including inheritance, fine granularity of permissions and changing permission during document lifecycle ...

Indexing Protected Content II

- Document Processing
 - Resolve ACLs to SIDs
 - Sent to Indexer with document
- Query Processing
 - Authentication by Front-End
 - creates filter with expanded user credentials and group memberships



Overview of Indexed documents

Documents indexed by CERN Search - March 2015		
CDS	1,766,190	
EDMS	4,927,599	
Indico	3,445,490	
Jacow	268,520	
Twiki	125,702	
Drupal	33,871	
Jira	14,613	
SharePoint (eSpace/Groups Archive)	14,754,487	
CERN Websites	1,570,301	
Total	~27 million	

 protected contents (SSO integration)

Upcoming Search Features

- CERN Search Migration: SP 2010 → SP 2013
- Find content
 - In-page documents preview
 - Visual Refiners

The screenshot displays the CERN Search interface. At the top, it says "CERN Accelerating science" and "Signed in as: bsilvade (CERN)". The search bar contains "esn". Below the search bar, there are tabs for "Everything", "People", "Conversations", and "Videos". The search results are filtered to "Preference for results in English".

The results list includes:

- sparglight-wp-1202-altimeter-esn**: By Charlene Li with Alan Webber and Jon ... Altimeter defines Enterprise Social Networking (ESN) as: A set of technologies that ... 2. But ESNs are not simply Facebook behind a firewall ... [social.cern.ch/me/bsilvade/.../sparglight-wp-1202-altimeter-esn.pdf](#)
- The Rise of Enterprise Social**: According to a July 2012, McKinsey & Company study, there's over a trillion dollars that could be unlocked by utilizing social technologies! ... Yet only 3% of the companies in the ... [social.cern.ch/me/psiatka/.../The-rise-of-enterprise-social.docx](#)
- Social Networking at CERN ITTF**: Bruno Silva de Sousa IT/OIS ... Please go to social.cern.ch and explore ... Video about #ESN ... McKinsey's research indicates that interaction workers typically spend 28% of ... ESN components ... [social.cern.ch/me/.../Social-Networking-at-CERN.pptx](#)
- OIS**: Use of technologies to create value by connecting members of an organization ... Like public social media (Facebook, Twitter ... ESN components ... ESN deployments are a medium-long term effort ... [social.cern.ch/me/psiatka/.../presentation/Share.pptx](#)
- Tasks**: [social.cern.ch/me/bsilvade/Documents/Work/Tasks.one](#)
- The-rise-of-enterprise-social**: The Rise of Enterprise Social Networks by, Mark Fidelman According to a July 2012, McKinsey & Company study, there's over a trillion dollars that could be unlocked by utilizing ... [social.cern.ch/me/psiatka/.../The-rise-of-enterprise-social.pdf](#)

On the right side, there is a preview of a presentation slide titled "Social Networking at CERN ITTF" by Bruno Silva de Sousa IT/OIS, dated 28th March 2014. The slide content includes "Take a look inside Social Networking at CERN ITTF Before we start Agenda Problem Enterprise Social Networking" and "Changed by Bruno Silva De Sousa on 20/03/2014 01:00 PM".

Summary

- **CERN Web Services**
 - heterogeneous infrastructure to address needs of different CERN user communities
- **CERN Social – Enterprise Social Networking**
 - new lightweight communication channel that encourages collaboration
 - a single place to stay up-to-date about CERN
- **CERN Search**
 - provides access to a wide range of (protected) information and CERN document repositories such as CDS, Indico, Twiki, SharePoint, JIRA and (EDMS).

Thank you for your attention!

Questions ?