Lotus software

IBM

Lotus Discovery Server

Wayne Boxall
Senior IT Specialist
Lotus Software

IBM Software Group

Agenda

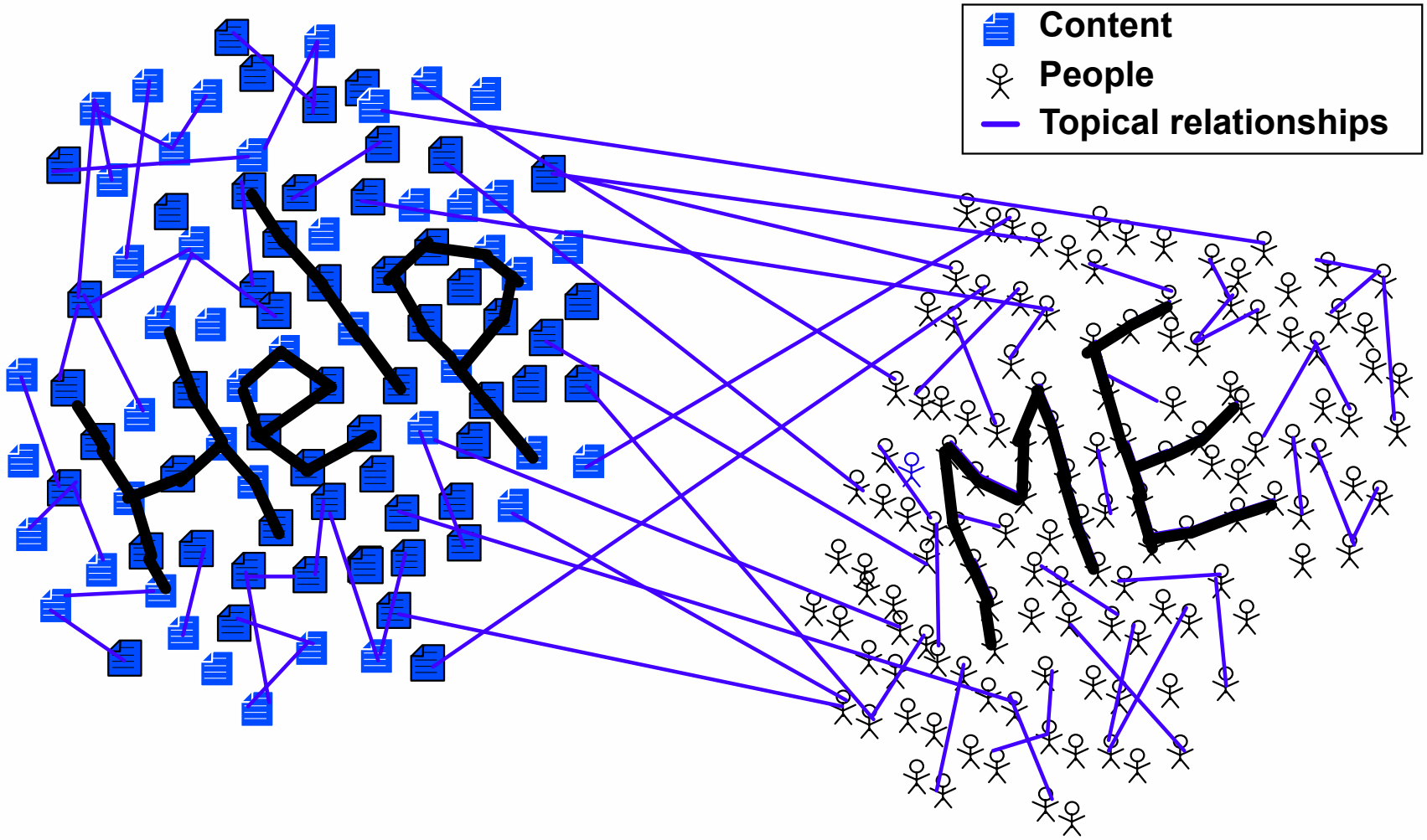
- Discovery Server Overview
- Architecture
- K-map (Corporate taxonomy) creation and modification
- Metrics Collections and Processing
- User Affinities
- API/SDK
- Futures

Lotus Discovery Server

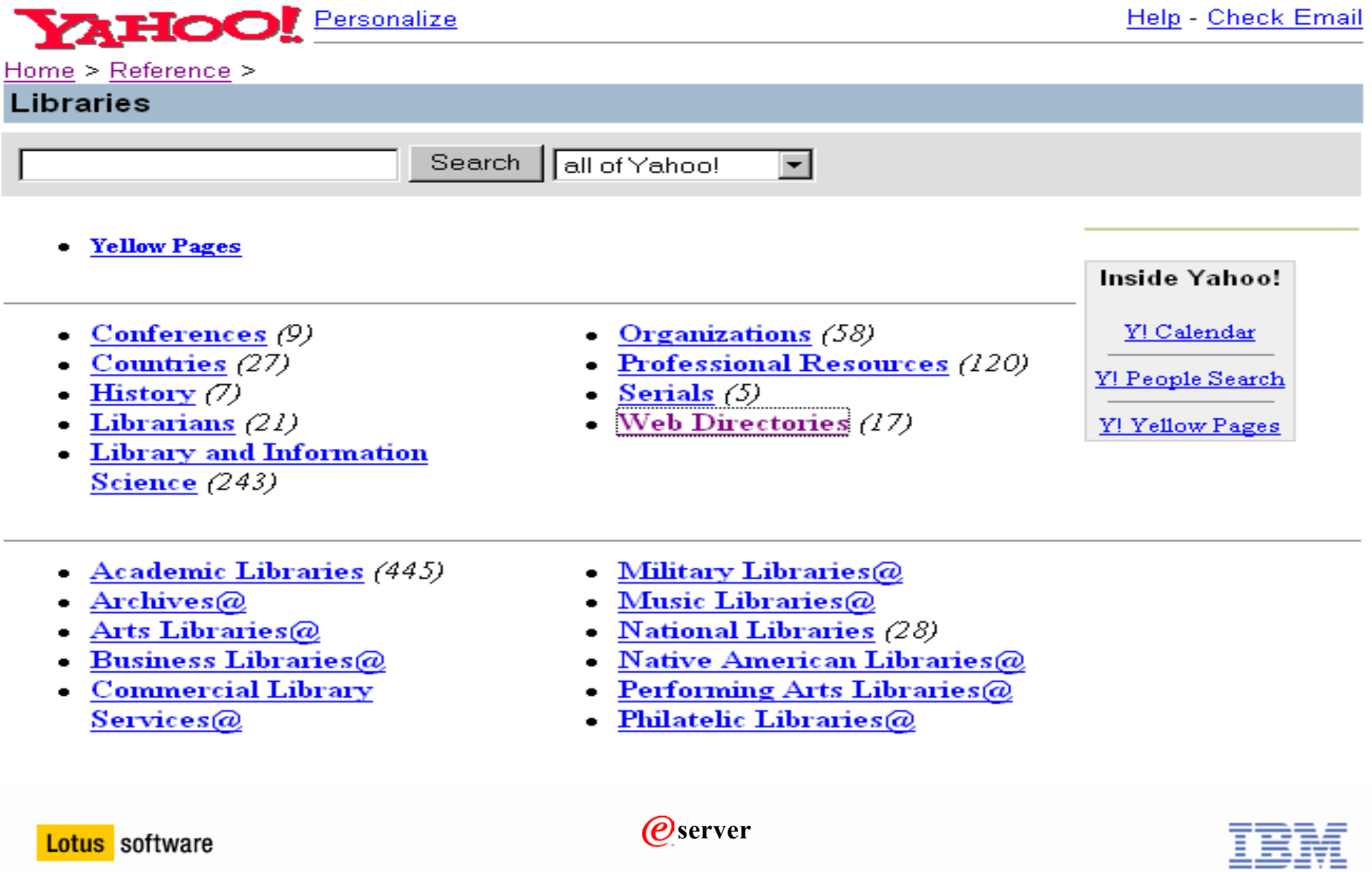
- The World's First Integrated Knowledge Server

- Connects People with the Right Info at the Right Time
- Integrates People and Documents into a single topical Knowledge Map
- Discovers Relationships between People, Content and Topics to add *Context* to Information
- All while respecting Privacy and enforcing Security

Relationships!



Public examples of a taxonomy



What Does the Discovery Server Do ?

- **Discovery Server will:**

- create a knowledge map
- generate affinities
- create expertise profiles
- assign content value
- index everything

- cluster and organises documents
- relationships b/t people and topics
- mine skills, locate experts
- based upon computed metrics
- search for docs, people, topics, etc.

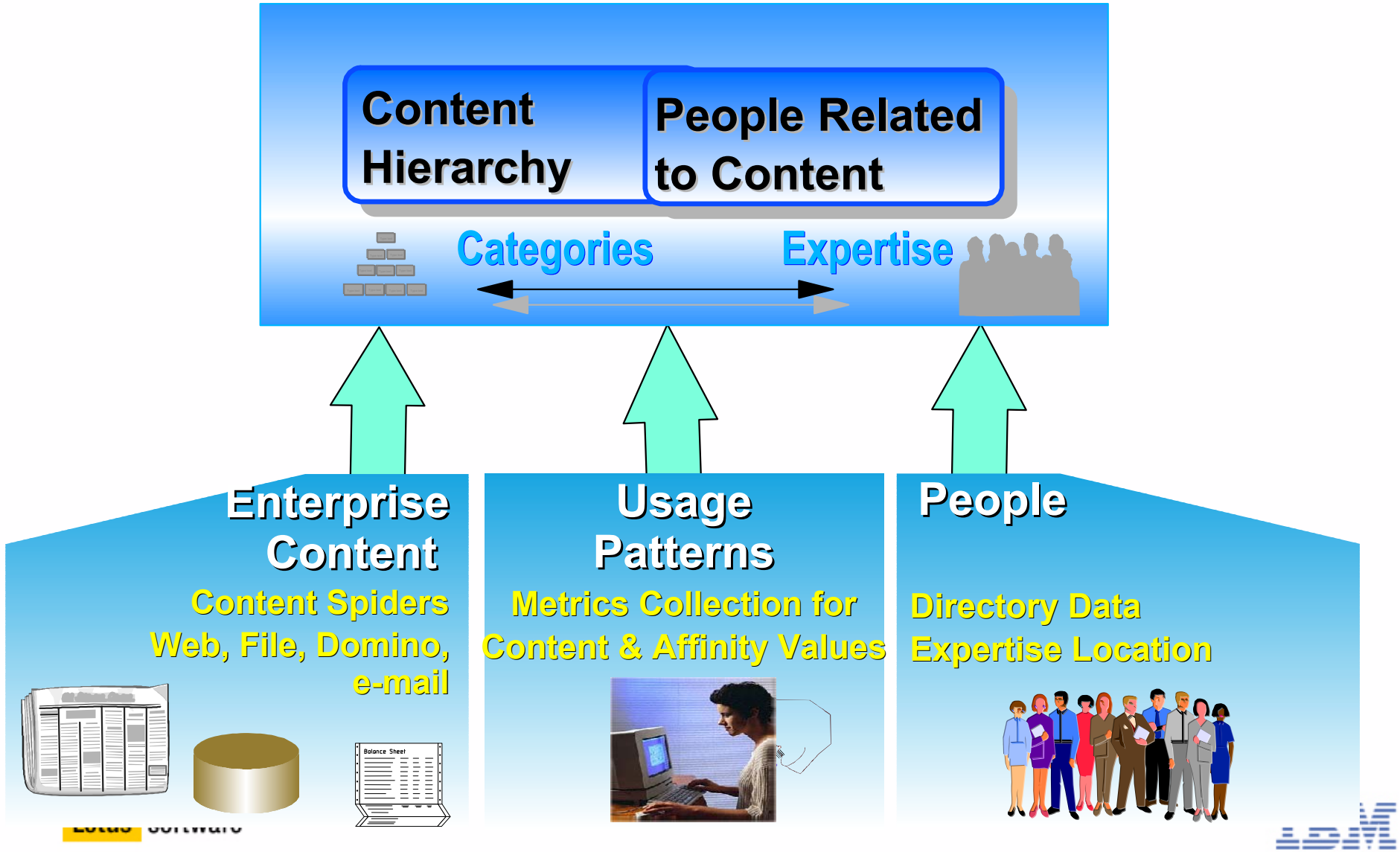
Discovery Server components constantly maintain and update themselves through a combination of automatic processes and administrative tools

Discovery Server Details

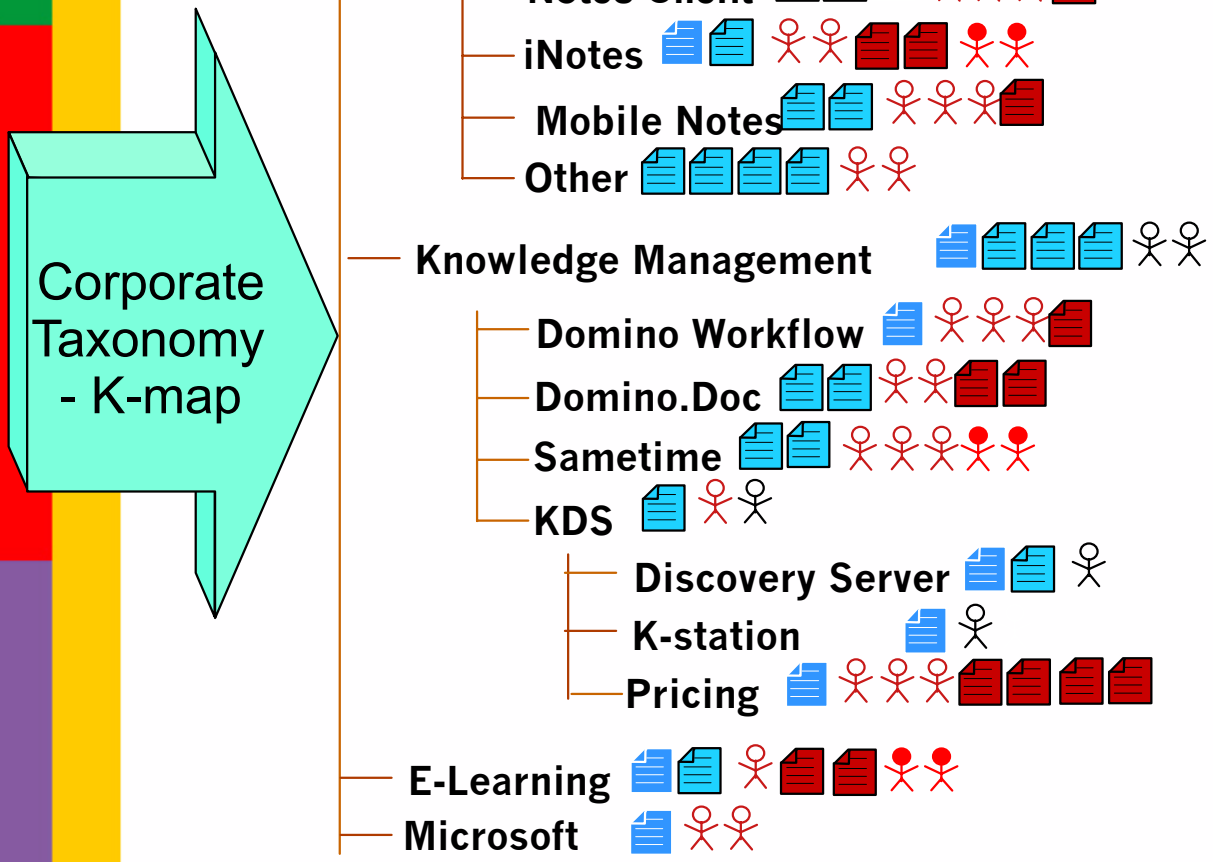
- Discovery Server must be on a separate server
- Translation and text processing coverage for international languages
- Client requirement - IE 5.01 SP2 or higher in R1
- Included Technology
  - Domino, Sametime
  - DB/2, IBM Sabio Clustering
  - InXight, KeyView...
- APIs:
  - Discovery Server Java API
  - Discovery Server XML Spider SDK (2.0)

DS Organizes Content Across Sources

Calculates Expertise and Information Values

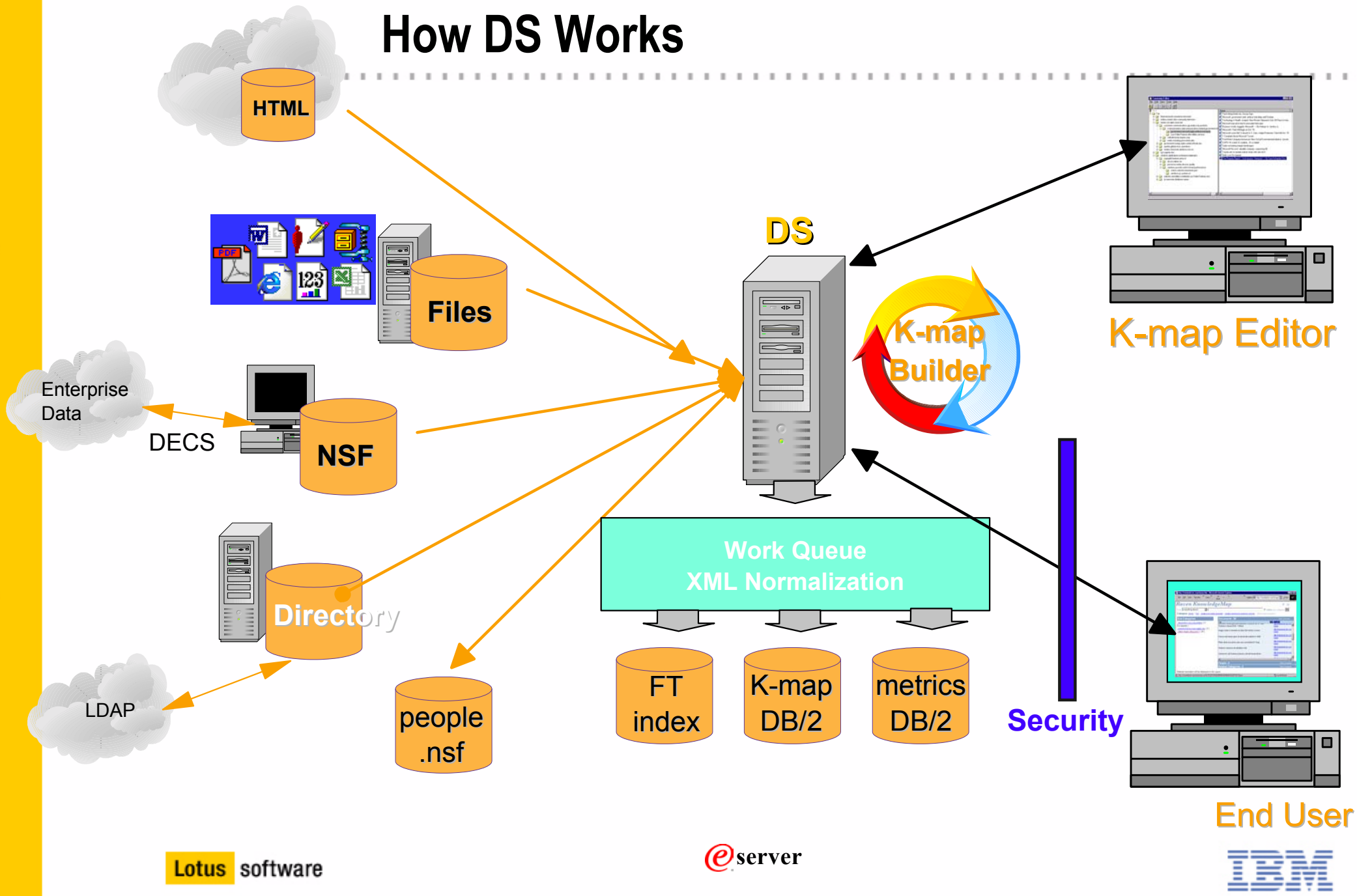


Discovery Server Services



- ***K-map (Taxonomy) Building***
  - Document clustering to categories
    - Labels categories
    - K-map Editing
  - Classification/Categorisation
- ***Expertise Profiling***
  - Synchronise Directory Data
  - Affinities/Expertise areas
    - Relationship to category/documents
  - Skills, Background
- ***Metrics Calculations***
  - Affinity Scores for Experts
  - Document Values
  - Information Utilisation Reports

How DS Works


HTML
Files
Enterprise Data
DECS
NSF
Directory
LDAP
people .nsf
DS
K-map Builder
K-map Editor
Work Queue
XML Normalization
FT index
K-map DB/2
metrics DB/2
Security
End User

Administration Overview

1. **Configure the server**
   - **Setup spiders and services**
2. **Generate profiles**
   - **Specify source of profiles - notes, LDAP etc**
3. **Create the K-map**
   - **Choose data repositories required**
4. **Refine the K-map using the K-map editor**
5. **Generate Affinities**
   - **Uses metrics - can also scan user's email**

All done via the Discovery Server Control Centre

Discovery Server Spiders

Content Spiders

- Domino
  - QuickPlace
  - Domino.Doc
- File Directory (NT)
- HTML/ Web

- Extract content, metadata and normalizes to XML:
  - Over 40 file types
  - Security (ACL) ,Location to sources (used by search)
  - Text (for word, clustering, category analysis)
  - Author, Editor, Responder, etc. and Usage data (used by Metrics)

Directory Spiders

- Domino NAB
- LDAP v2 v3

"Base" person data (Name, Address, E-mail, etc.)

E-mail Spiders

- Domino Mail
- MS Exchange Under Dev

- Option to analyze approved by end user
- Analyzes content that clusters with existing K-map categories,
- Matches augment expertise scores
- No data source pointers stored, no public access to e-mail document

Other Content

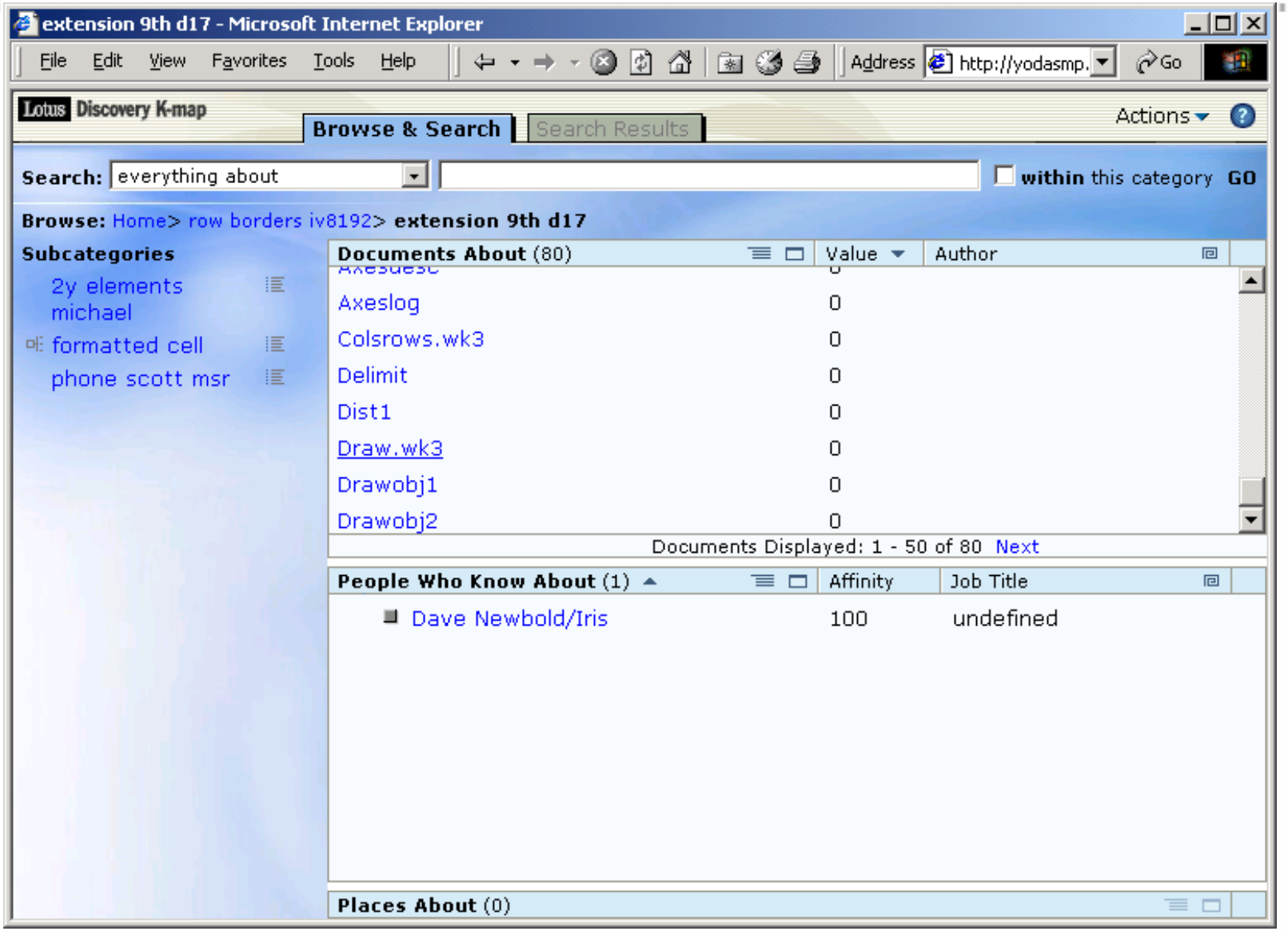
- LDS Domino Spider
  - DECS Connectors
    - RDBMS / ERP
- DS XML Spider (SDK)

About the Knowledge Map (K-map)

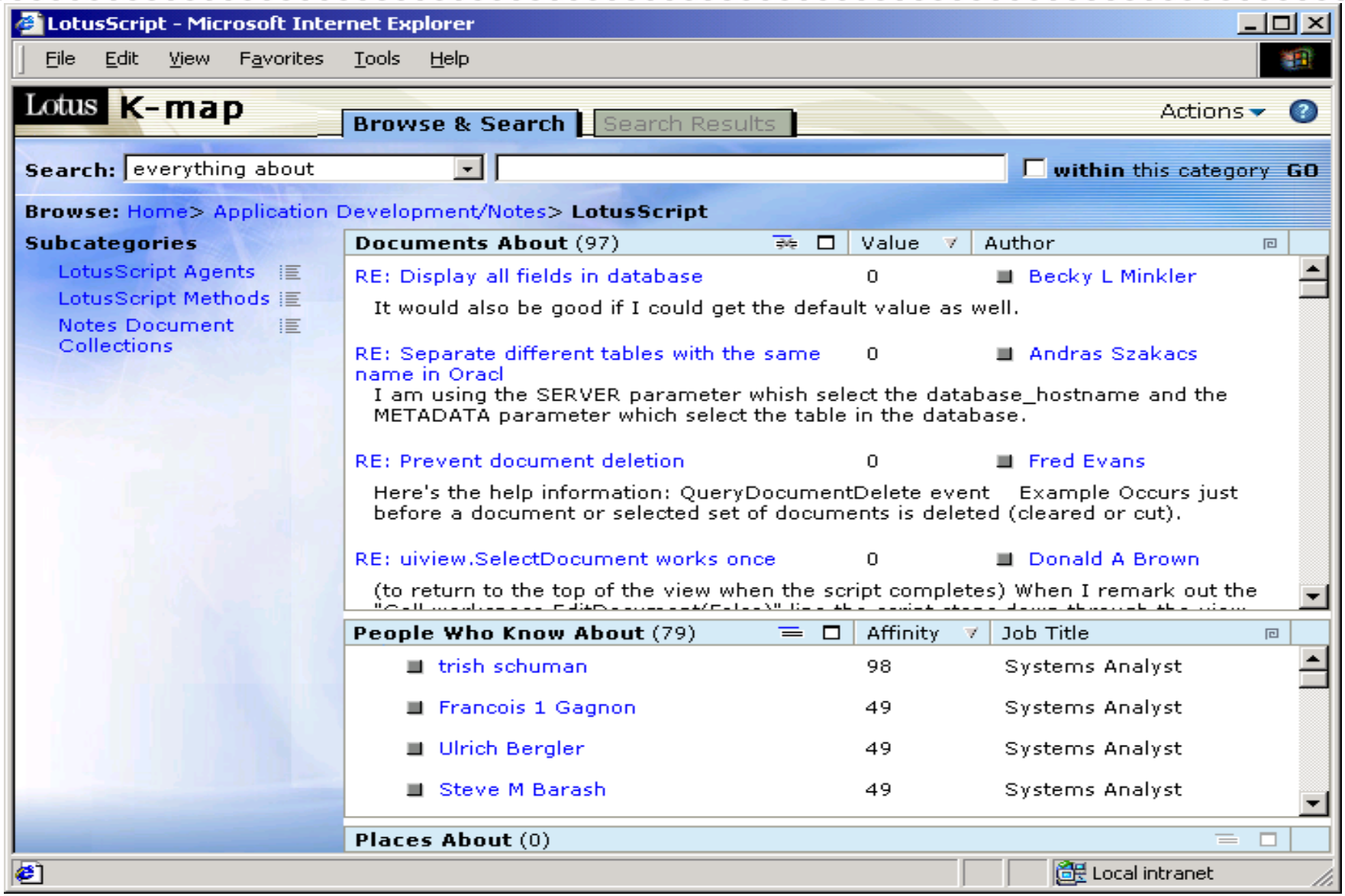
- K-map is for browsing by your end users, identifying names of affinities, and training for categorisation.
- Users drill-down through category tree until they find a category that could contain a document they need.
- Designed for context sensitive searching

Human readable tags of clusters of documents

Garbage in = garbage out



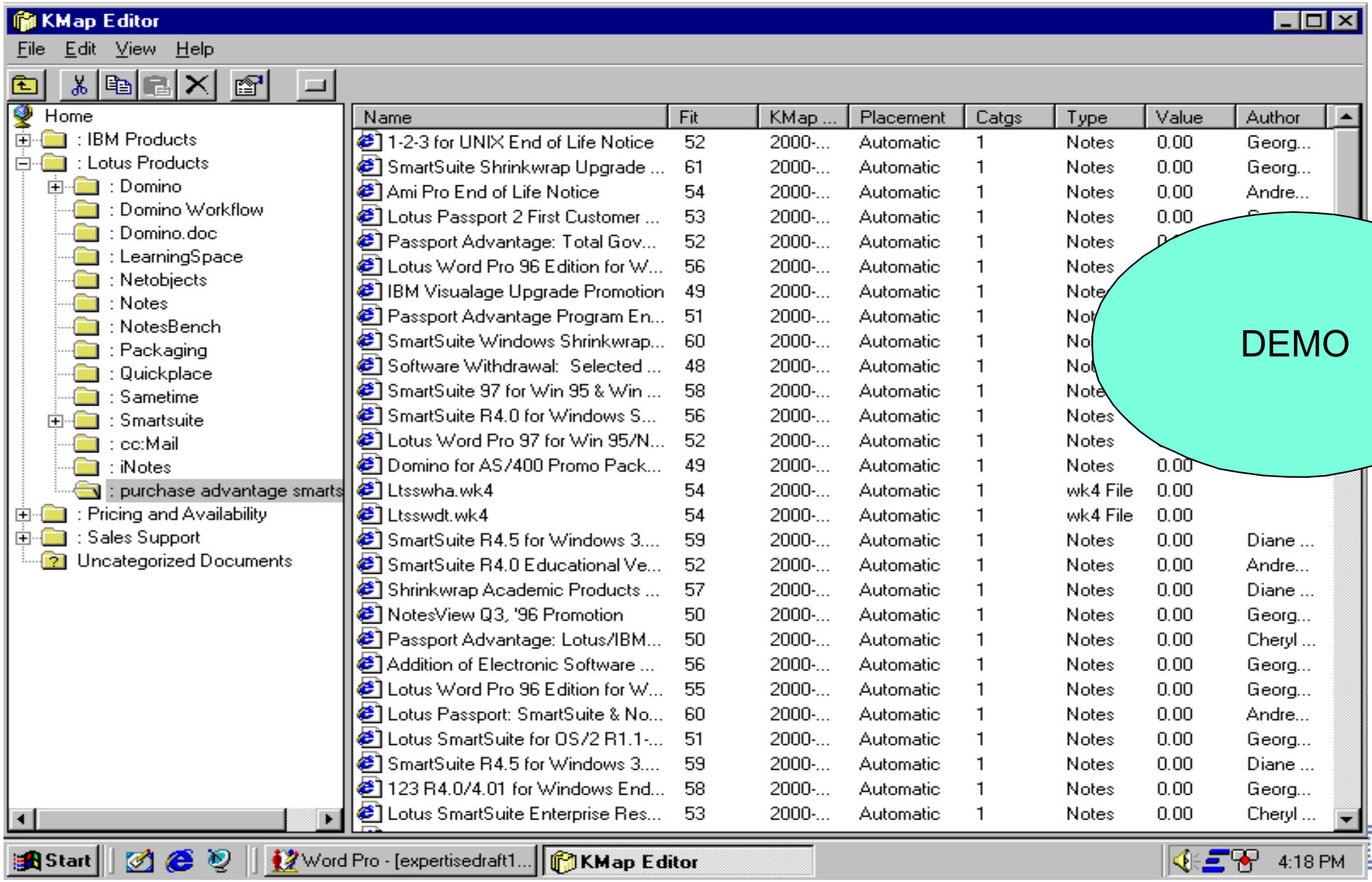
Richer results



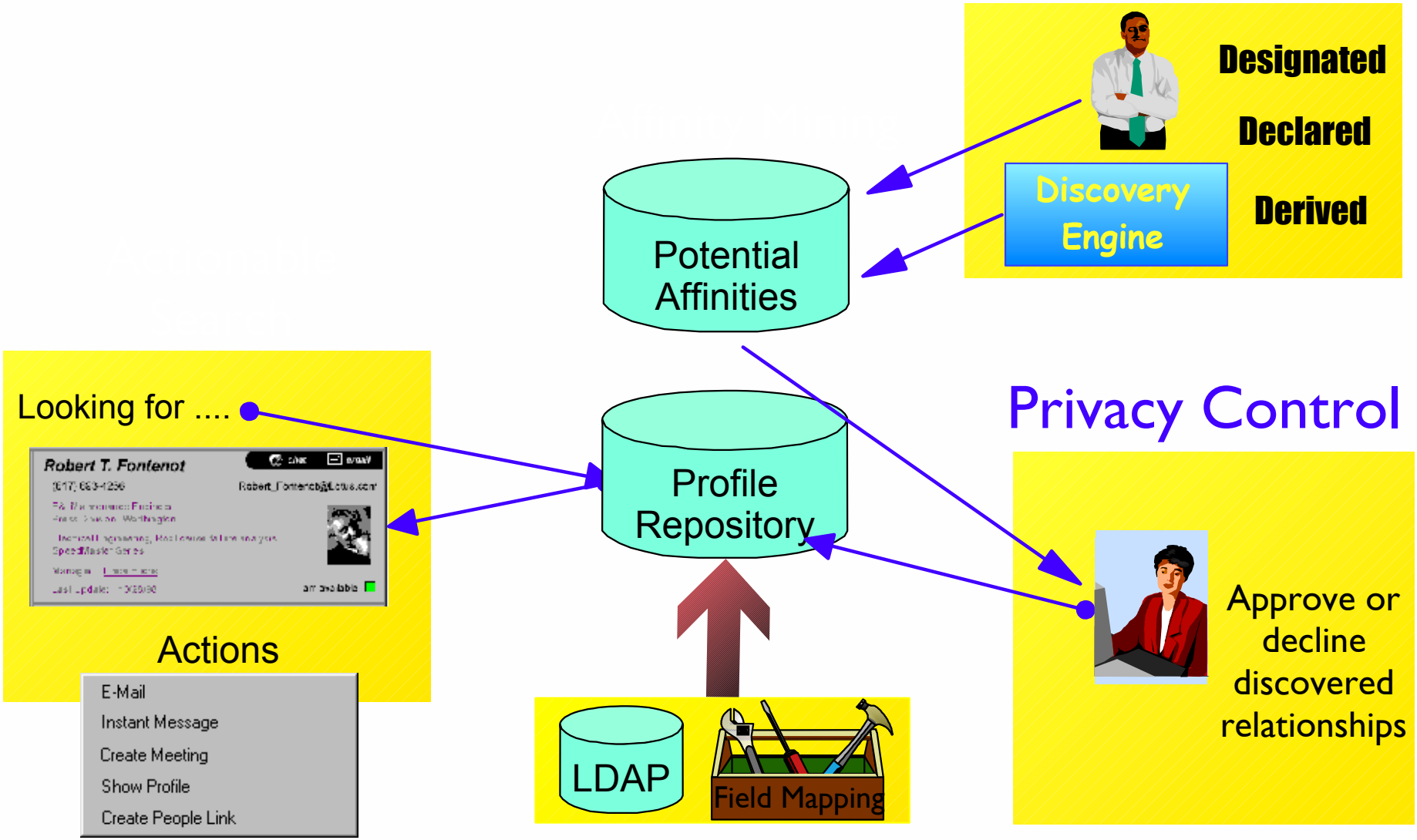
Time For Human Intervention!

- These initial clusters will almost always need a human editor, to re-label and re-categorise to make the K-map meaningful to the organisation.
- Humans have to inte[illegible]nd refining rules, but by analysing the c[illegible] ensure they are meaningful to thei[illegible]
- Remember: The K-[illegible]n by example
- Once the K-map is reorganised, new documents spidered move to the new structure. Old documents stay where they are.

K-map Editor View: Clustered Documents



Profile Creation and Affinities



Affinities and Document Values

- Powered by Metrics - the "Unsung Hero"
  - "Digital Breadcrumbs"
  - Statistics about information flow
  - No additional burden on users
  - Leverages document meta data
    - HINT: start getting your meta data in order now
  - Analysis reveals trends, relationships, and patterns

LDS Metrics Calculations

- Document Metrics
  - Links To
  - Links From
  - Responses
  - Opened
  - Recency of Accesses
  - Relevance

✓ ***Document Value***

- People Usage/Category Metrics
  - Author
  - Edit
  - Link To
  - Responses
  - Reading via Kmap

✓ ***Affinity/Category Rating***

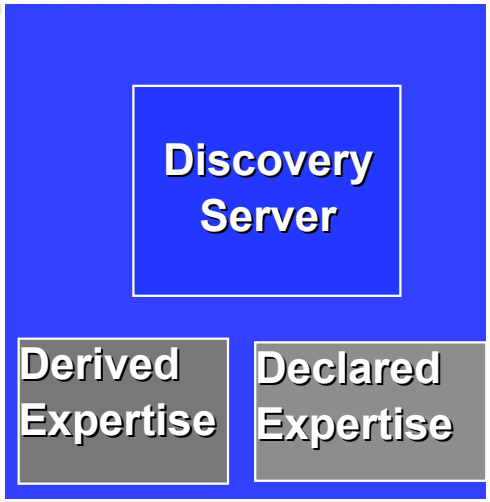
Automatic Calculation of User Affinities

- Metrics Collects User Interactions to Documents in Categories

- Affinity Threshold Set for Organisation (Administrator)
  - Is % affiliation to calculate prior to affinity notification
    - Usually 50 - 60%

- Metrics Calculations and Update Task...
  - Affinity Calculated for user/category = Threshold?
    - Yes - Propose Discovered Affinity to User (e-mail)
    - No - Continue to aggregate affinity data

  - Approved Affinity to Category displays in the profile and K-map search result



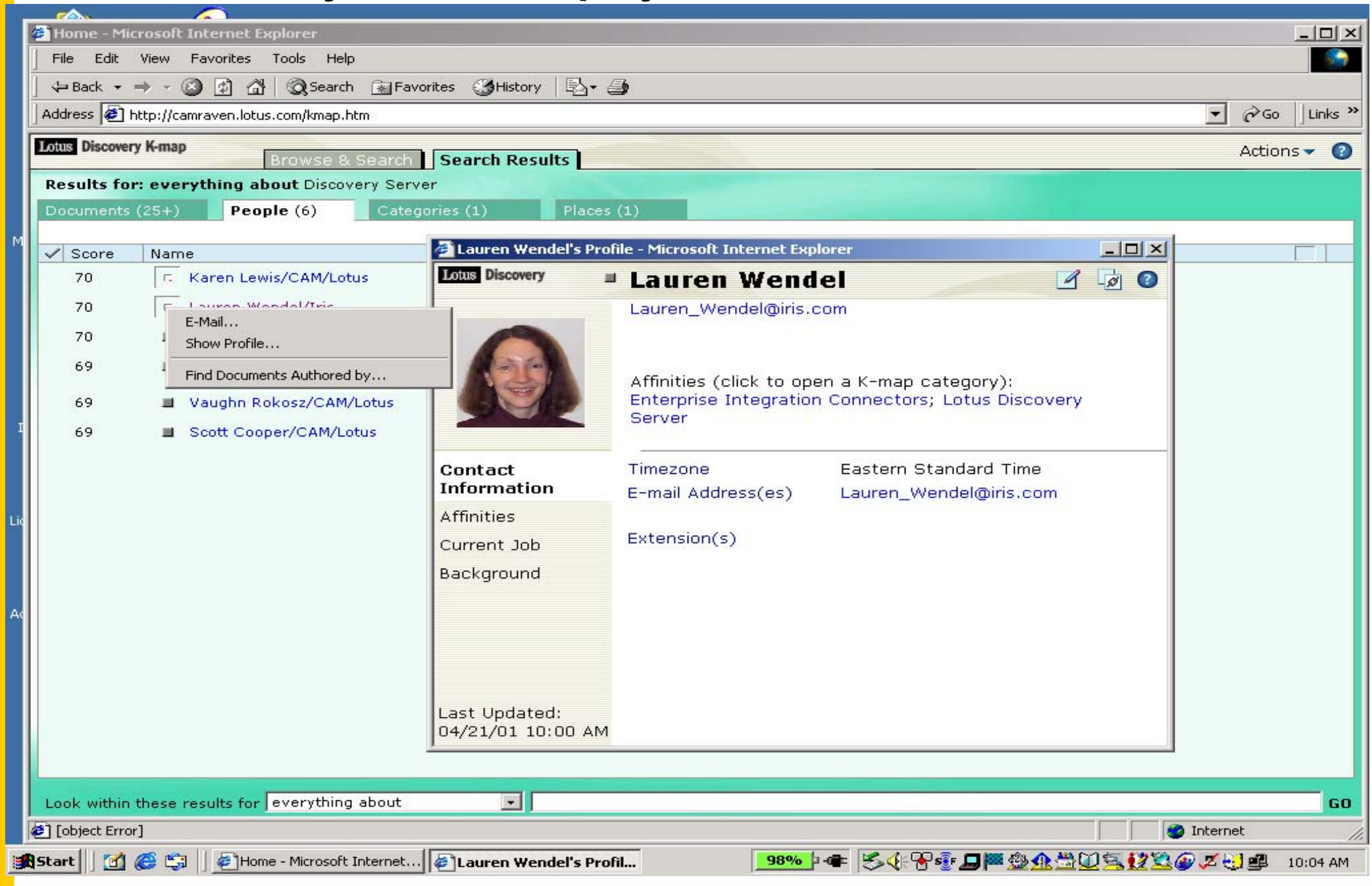
People Who Know About - 3 Sho

✓	Name	Affinity
	Thomas Libretto	96
	Chris Crummey	90

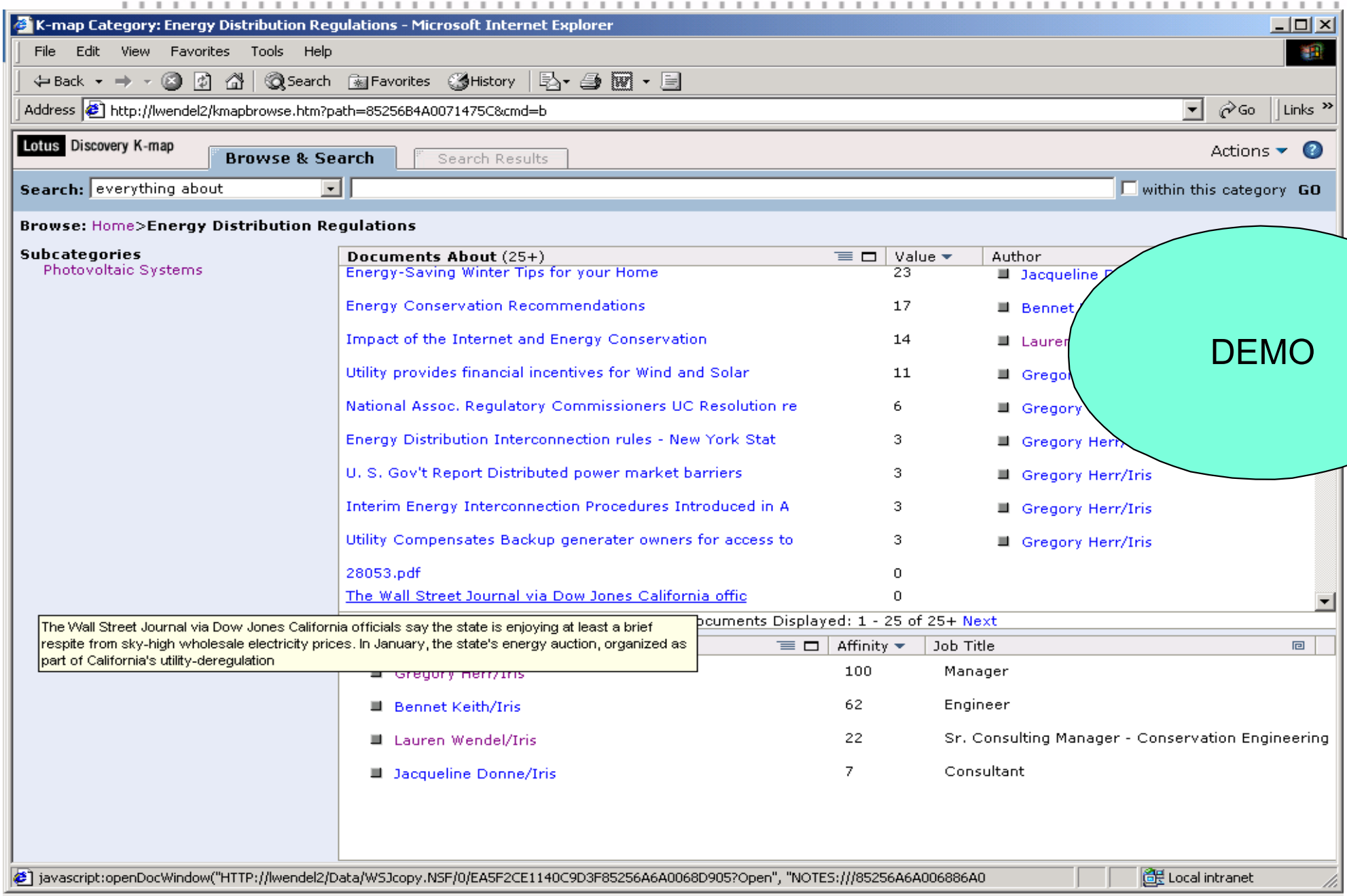
Lotus software

eserver

Affinity Values Displayed



Information Values in Context



Metrics Reporting: Data Usage Patterns and Analysis

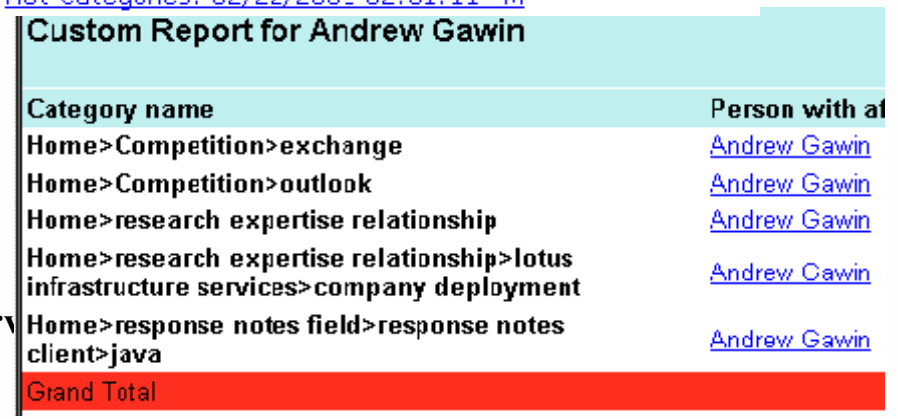
- Metrics Reports (Admin access)
  - ►Standard
  - ►Customised

Content Area with greatest activity?

Hot Categories

Content Area with most experts?

Categories with affinities

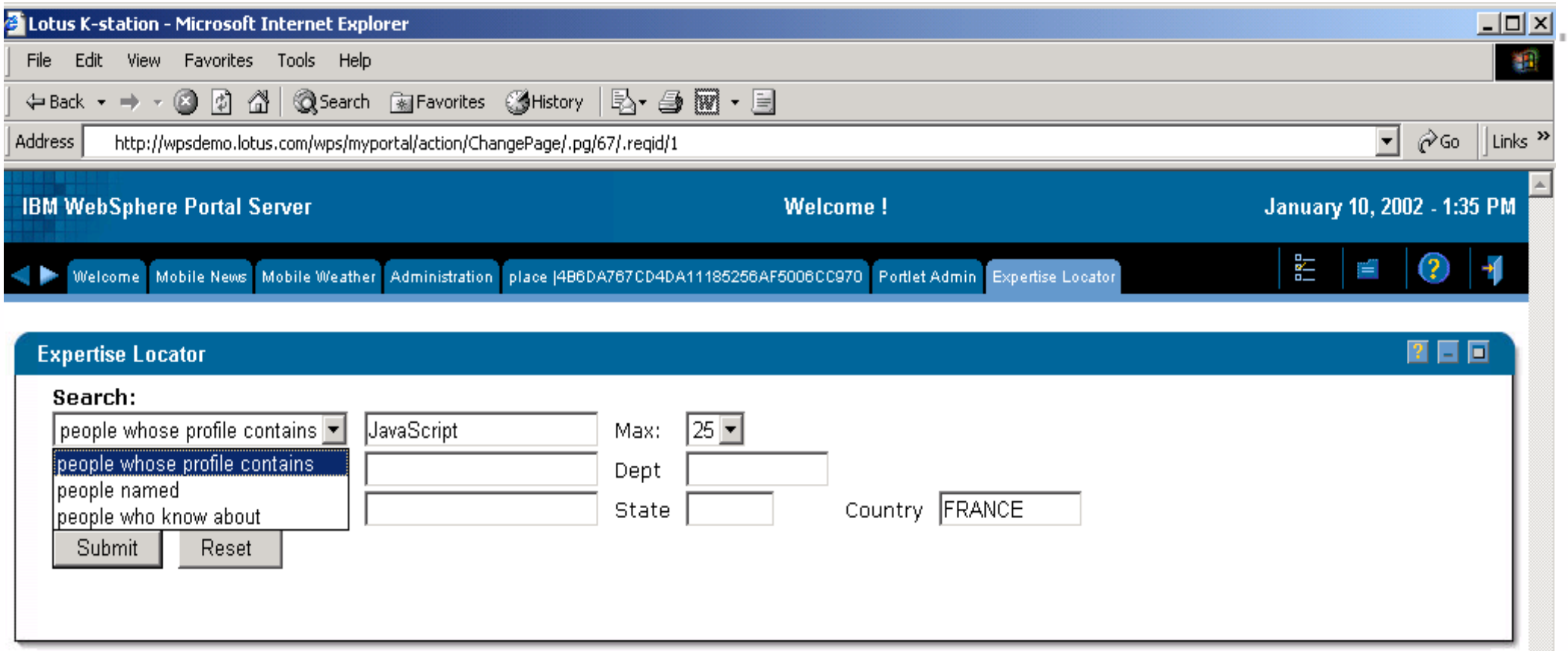


Lotus software

Java API

- New in 1.1
- Integrate LDS services into your applications
  - Customise display, interaction with processes and data
- Track submitted searches
- Recommendations and Referrals
- Custom UI for K-map editing or corporate portal
- Registration of content for spidering
- Tampering with Categories

Alternate Search Interface using DS APIs



XML Spider SDK

- In DS 1.2/2.0
- XML Specification
- Access additional content sources
- Promoted to LDS
  - Proprietary repositories
  - File System which are not Win32 compatible
  - Databases not supported by DECS

Pocket Discovery Server

"Pocket LDS" Mobile Access to Customised Subscriptions of Discovery Server Content and Experts

Futures

1.1 Shipped Oct 2001

- Stand alone server from K-station
- Bug fixes
- Includes Discovery Server Java API Toolkit 1.1
  - - Customize display, interaction with services and data

1.2/2.0 to ship H1 2002 - some of the features

- Exchange Spider, XML spider toolkit, Active Dir Spider
- Improved Performance – No Java applet
- API Enhancements
- Taxonomy Import Ability
- View and Selection Formula for Domino Data selection
- Improved UI and KMap Editor

Additional Product Information

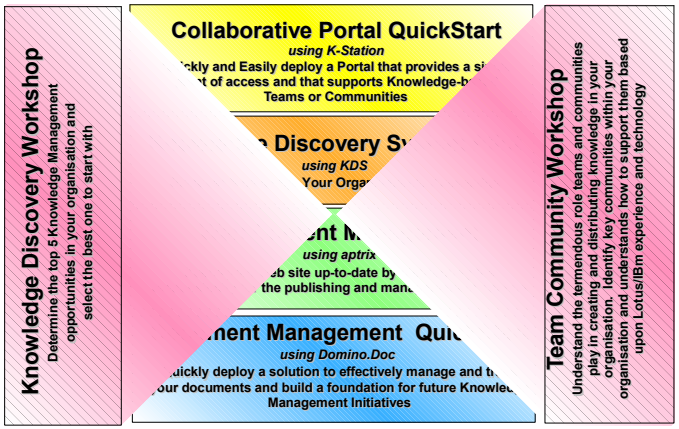
- www.lotus.com/KM
- www.lotus.com/discovery
- www.notes.net - Iris Cafe KDS Discussion
  - Pocket LDS

- e-mail: Wayne Boxall
  boxallw@au1.ibm.com

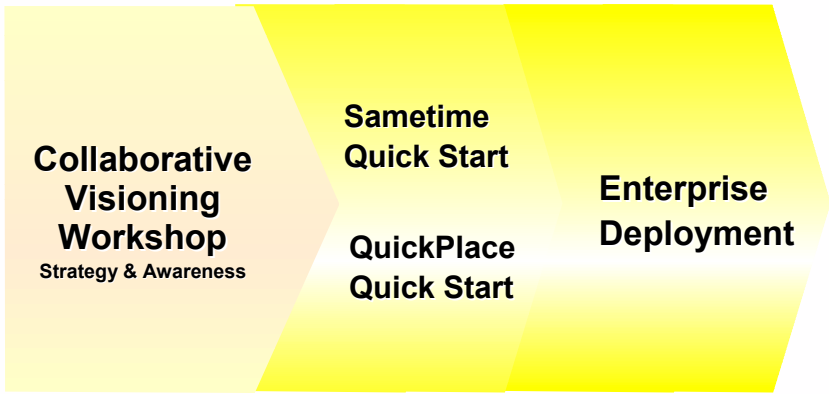
Offerings Portfolio

IBM Software Group Services

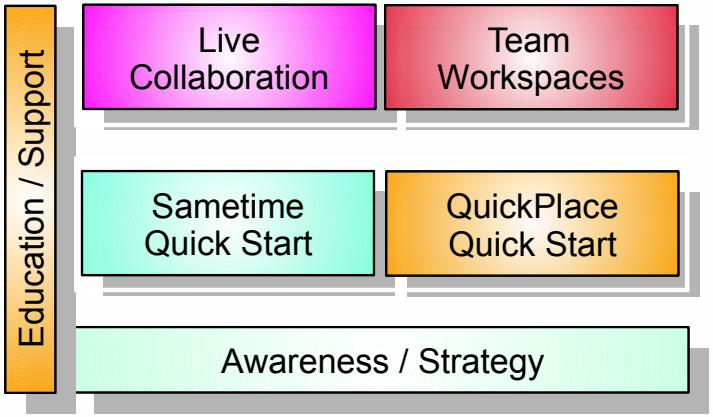
Advanced Collaboration



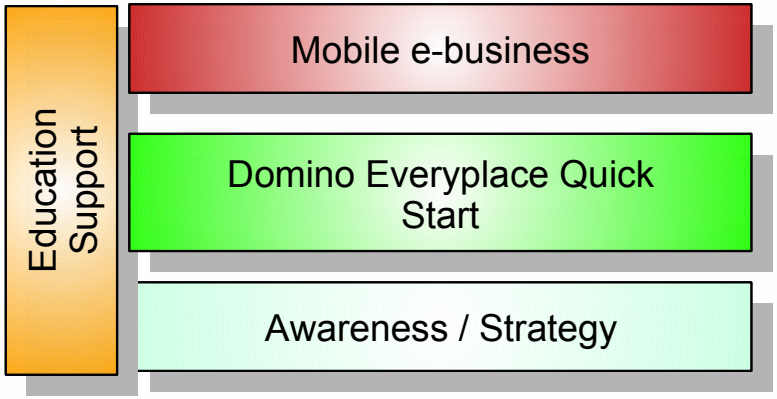
Business Collaboration



c-Commerce



Mobile e-business



The Lotus Knowledge Management Portfolio...

Knowledge Discovery Workshop

Determine the top 5 Knowledge Management opportunities in your organisation and select the best one to start with

Collaborative Portal QuickStart

Quickly and Easily deploy a Portal that provides a single point of access and that supports Knowledge-based Teams or Communities

Lotus Discovery Server Pilot

using LDS

Discover & Leverage Your Organisation's Knowledge

Web Content Management

using aptrix

Keeps your web site up-to-date by distributing the responsibility for the publishing and management of content

Document Management QuickStart

using Domino.Doc

Quickly deploy a solution to effectively manage and track your documents and build a foundation for future Knowledge Management Initiatives

Team Community Workshop

Understand the termendous role teams and communities play in creating and distributing knowledge in your organisation. Identify key communities within your organisation and understands how to support them based upon Lotus/IBm experience and technology

Questions?

Thank You

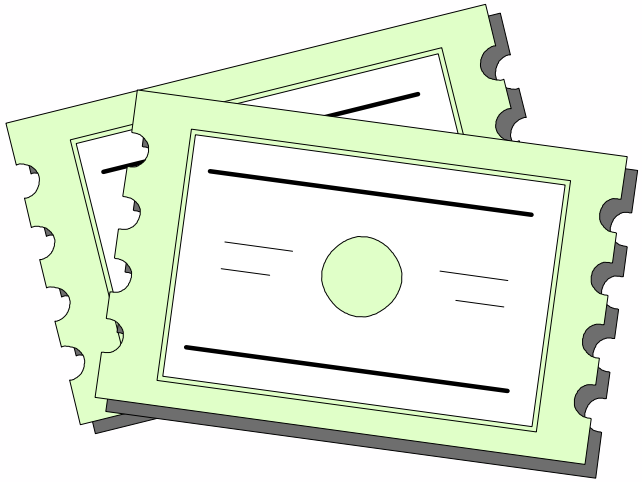
Lotus software

eserver

IBM

REMINDERS!

- SWG Services for Lotus Exhibition Stand
- Visit the eServer team
- Business Partner Showcase
- EVALUATION forms



Lotus software

Lotus eSolutions Agenda

8.00am Registration and Breakfast

8.30am Lotus Welcome: Duncan Hewett, GM, Lotus Software, ANZ

8.45am Lotus Software Keynote: Scott Cooper, Sr. VP Lotus Software WW

	Technical Solutions	Business Solutions
9:30am	Lotus Discovery Server	Australian Bureau of Statistics: KM Strategies
10:15am	Morning Tea	
10:45am	Lotus Technology co-existing with your current infrastructure	Knowledge Management Roadmap: Lotus Discovery Server, Domino. Doc and Lotus Workflow
11:45am	WebSphere Portal Solution	e-Learning Solutions: IBM Mindspan
12:30pm	Lunch	
1:15pm	eCollaboration: A technical "under the covers" to integrating your desktop	eCollaboration: iSeries and Domino 6
2pm	Afternoon Tea	
2:30pm	Lotus Notes, Domino 6 & Designer: Technical Overview	Next Generation: eWorkplace
3:30pm	Lotus Sametime & QuickPlace: Developing on the Sametime Platform for real-time collaboration	Content Management Framework: Putting Domino on a Diet

Lotus software

IBM

Lotus Discovery Server

Wayne Boxall
Senior IT Specialist
Lotus Software

IBM Software Group

Building the K-map
Start With Your Richest Documents

- Categories with lots of documents that use the same unique words to be better able to attract similar documents
- Generic Notes databases with rich meta-data and preferably long, rich-text.
- Domino.Doc databases. Choose the "Open attachment's container" option on the Data Repository form.
- Not too many documents! Better to have say 3K...so you have time to check the original clustering.

Start With Your Richest Documents (cont.)

- Windows-compatible file system files with rich meta-data.
  - If using a Word document, for best results make sure that the properties fields contain up-to-date Author and Title information.
  - Presentation files are good if their meta-data is good and if the slides are very wordy.
  - Spreadsheet files are not recommended.

  ....if the author info is not filled out these documents will not contribute to generation of affinities.

Start With Your Richest Documents (cont.)

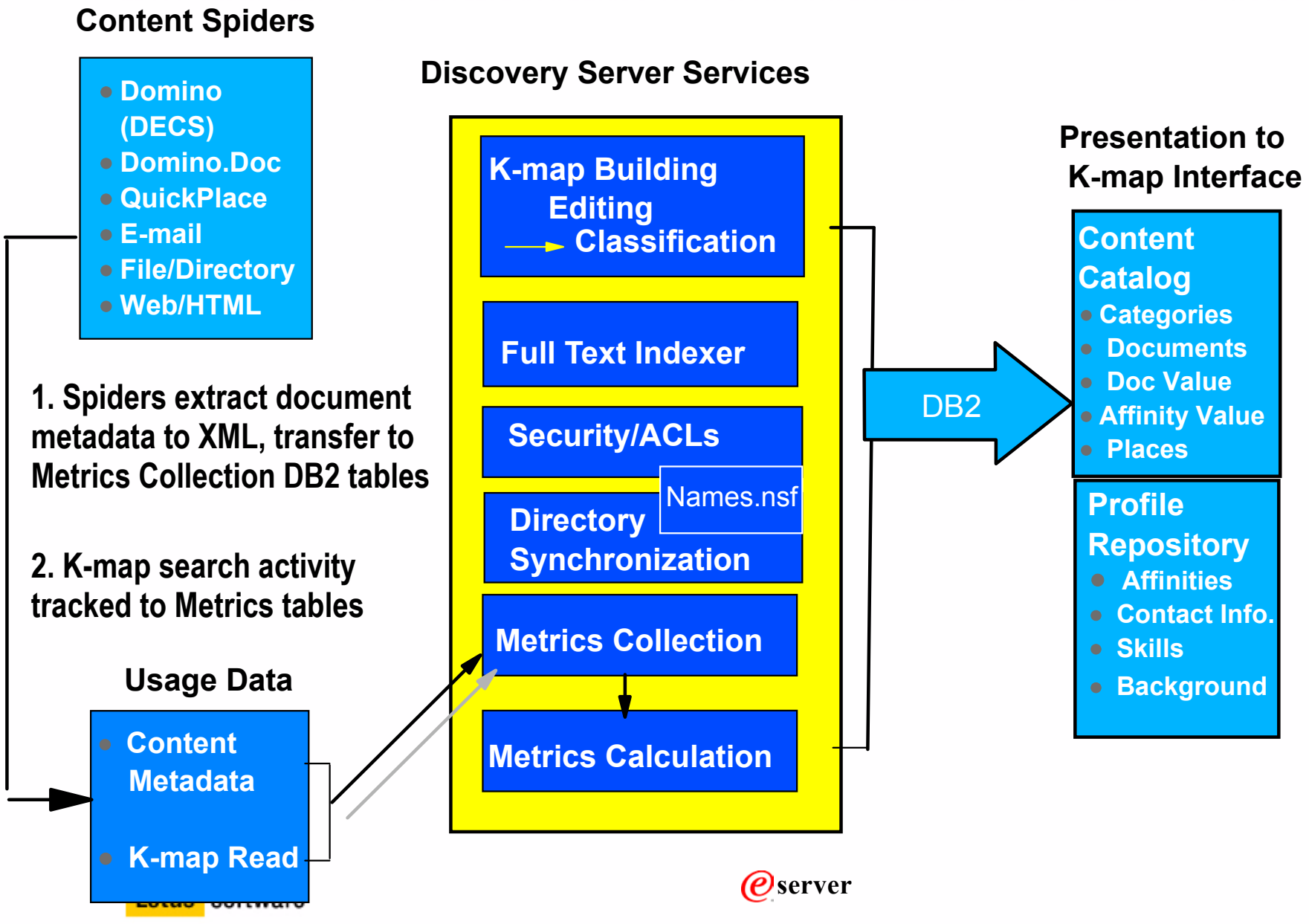
- Quickplace databases, if the documents contain lots of text. Choose the "Open attachment's container" option on the Data Repository form.
- Web sites with good meta-data and lots of text. If in doubt about a Web site's meta-data, right-click one or more pages on the site and choose View Source.
- Must have enough tokens >30 (words) to make the content worthwhile and to have it included in the K-map.

How the K-map builder works

- Operates in 2 modes: Clustering and Categorization
- Uses vector analysis of the words in documents to group similar documents into clusters
- Uses a combination of Expectation-Maximization and K-means to build these clusters
- Based on SABIO software developed by IBM
- Treats words within documents as points in a large multi-dimensional space.
- Teases out general themes in collections of documents
- Will require manual reorganisation

Metrics Collection

Captures Usage Data from Content and K-map interactions



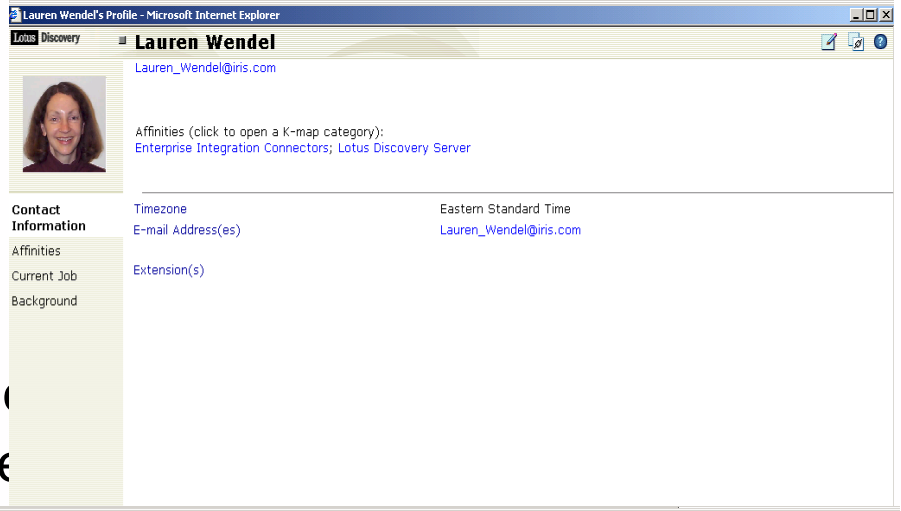
eserver
IBM

LDS Affinities

- Derived - LDS 'discovers' through metrics calculations affinity to category strength
  - ►proposes to profiled user for approval and publication to profile
- Designated - Other user (with designation privileges) requests category published to profile
- Declared - User 'self declares' affinity directly via K-map category selection

LDS Expertise Profile Data Model

- Centrally Organises:
  - ► Directory & Supplemental Data
    - – Synchronises updates
  - ► Background, skills,education data
  - ► Affinity (Expertise Areas)
    - – Affiliations to K-map category d
    - – Statistically computed to all othe
    - – Discovered, Declared and
- Displays:
  - ► Approved Affinities
  - ► Contact Information, Backgrou
  - ► Sametime Integration
- Privacy Controls:
  - ► End User can edit Profile
  - ► User controls publication of t
    others



Lotus software

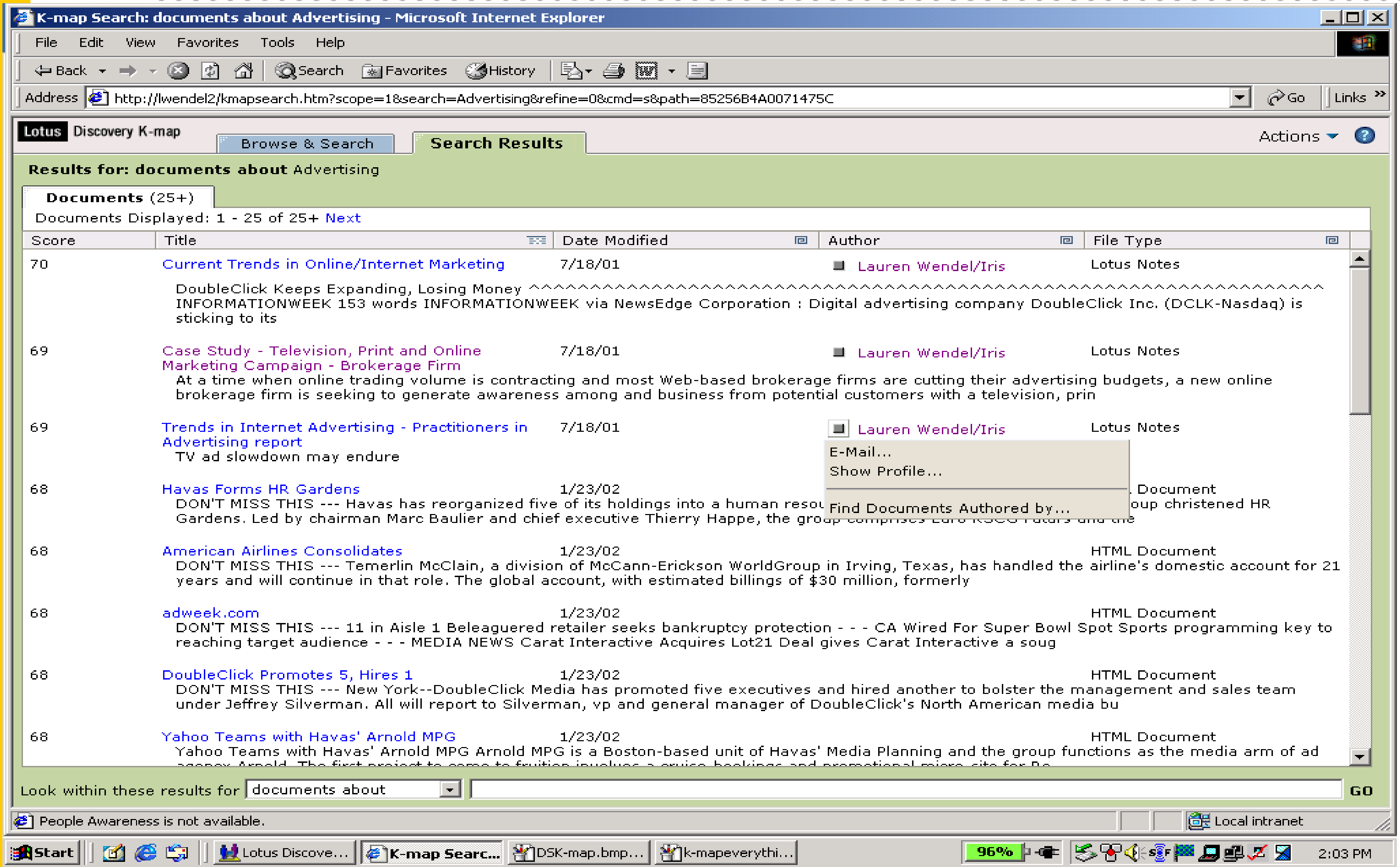
Automatic Calculation of Document Values

- **Usage Metrics assigned default weights:**
  - **Documents have higher values if Linked to, Responded to, Read...**
    - **Default weightings may be tuned by administrator**

- Metrics Calculations - **Calculates document values from Metrics Collections data**

- K-map Document Search result Value combines:
  - **Information Topic Search rank + Usage Metrics Score per document**

Information Values in Context

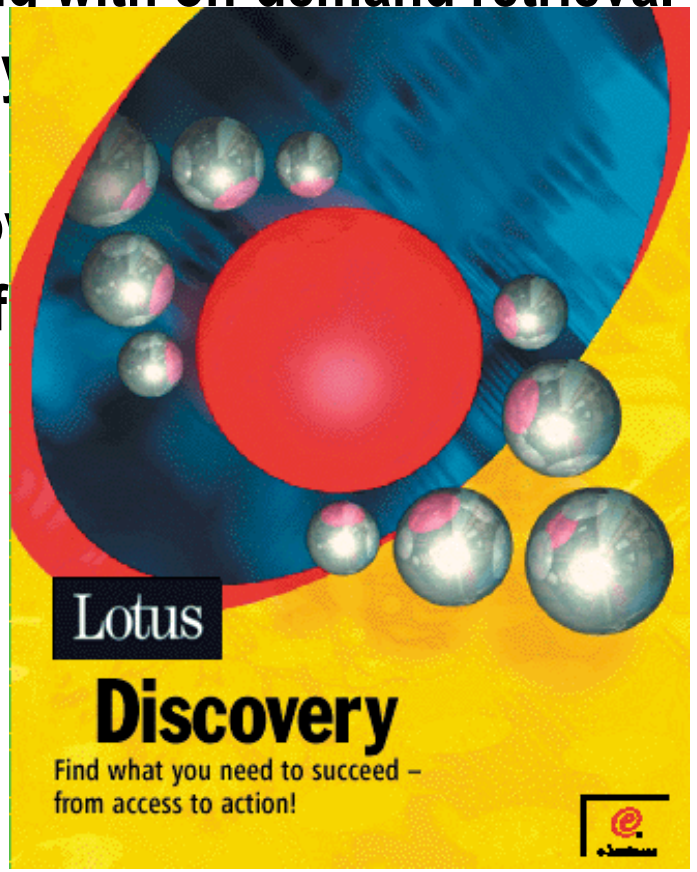


Search for Anything

- Discovery Server makes Search useful
- Index almost any content, almost anywhere
- Scope your searches, find only what you need
  - Everything About
  - Documents About
  - Documents Authored By
  - People Named
  - People Who Know About
  - People Whose Profile Contains
  - Categories About
- Score combines search relevance with metrics value

Search for Anything

- Discovery Server makes Search useful
- Index almost any content, almost anywhere
- Scope your searches, find only what you need
- Score combines search relevance with metrics value

Atomica for Lotus Discovery Server

- Combines intelligent data gathering with on-demand retrieval
- Makes LDS-based content instantly
  any application
- Quick single-click access to Disco
- Integrates structured information f
  systems and 3rd party content
- www.atomica.com