

# Supporting enterprise knowledge management with weblogs

A weblog services roadmap

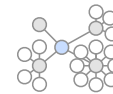
**Michael Angeles**



[jibbajabba@mailblocks.com](mailto:jibbajabba@mailblocks.com)



<http://studioid.com>



1

Good morning.

Today we're going to be talking about weblogs in the corporate intranet as tools for personal knowledge management, and how we, as the library, might help the organization use weblog output for organizational KM.

## Overview



**Roadmap** for enterprise weblog/k-log information services



**Different approaches** to blog services & **success factors**



Choosing **your solution**

2

We're going to be talking about the problems blogs present in terms of making their output findable.

What we're seeing with weblogs is a focus on **bottom-up knowledge creation**.

These grass-roots efforts are beginning in some cases to capture tacit knowledge -- the stuff that usually gets extracted in conversations.

But they present a problem in that they introduce distributed data that is often not collected and indexed for redistribution as an aggregate.

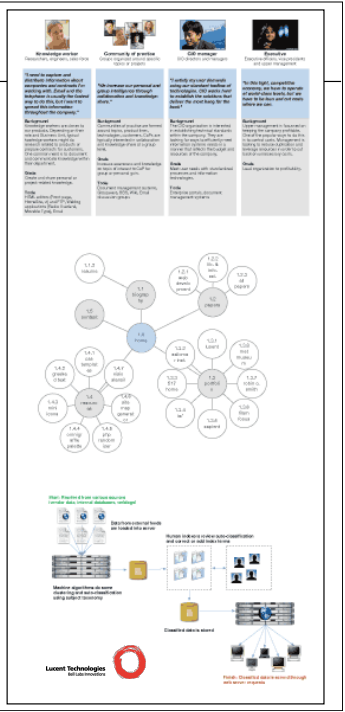
**The roadmap** I'll present is really a discussion starter for information services related to blogs. We'll look at some options to consider with the assumption being that we want weblog output to surface in the right places -- to prevent the invisibility of captured knowledge.

We'll also talk about connecting blog output to other types of enterprise information.

# About me

- Information Specialist, Lucent Technologies: Library systems design and development
- An information architect;  
Contributor to Asilomar Institute for Information Architecture;  
Contributor to Drupal open source software development
- On the web: [iaslash.org](http://iaslash.org),  
[urlgreyhot.com](http://urlgreyhot.com)
- In the intranet: Library staff blog,  
Library staff wiki

- # About me
- Information Specialist, Lucent Technologies: Library systems design and development
  - An information architect;  
Contributor to Asilomar Institute for Information Architecture;  
Contributor to Drupal open source software development
  - On the web: [iaslash.org](http://iaslash.org),  
[urlgreyhot.com](http://urlgreyhot.com)
  - In the intranet: Library staff blog,  
Library staff wiki



I work for an integration group in the library that works between the various clients, content owners, researchers and systems group.

I provide the information architecture, interaction design, and site development guidance on our projects.

This talk documents some of the investigation I'm doing to position the library with regard to blogs in the enterprise.

## About my organization

- Integrated Information Solutions: The digital library for **Lucent Technologies** (predominantly electronic services still circulate books by interoffice mail)
- Traditionally supported Bell Labs Researchers, now predominantly business users

**Lucent Technologies**  
Bell Labs Innovations



4

My organization provides the physical and digital library services for Lucent. We're a Bell Laboratories organization and traditionally supported researchers, but now provide services predominantly to business users.

### **About you. Some assumptions.**

- You know what weblogs are
  - A web site (usu. of personal/non-commercial origin) that is frequently updated with information and links to resources within a particular subject area.
- You've heard of a few weblog sites (e.g. Slashdot, BoingBoing, Rogue Librarian, Shifted Librarian)
- You know of a few popular weblogging applications (e.g. Movable Type, Radio Userland)

5

### **A few things I assume you know coming here today...**

You know what weblogs are. Basically they're a form of personal or community web publishing usually in journal-writing style.

You've heard of a few. Some popular ones are listed here.

And you know of a few applications you can use to publish a weblog.

## Basic concepts before we begin

- **Enterprise blog creation services**
  - Providing software or services for individuals create/publish weblog content
- **XML feed services for databases**
  - Any data can be made into RSS, some into RDF including news articles, press releases, email discussions, documents
  - A good method of providing access is to embed a link for an XML feed attached to database search results

6

There are **a few concepts** I refer to that we should have a common understanding of.

When I refer to **enterprise blog creation services**, I'm talking about providing a way for your users to publish their weblog with some centralized software controlled by you or your CIO.

When I talk about **XML feed services for databases**, I'm talking about providing **RSS feeds** for any type of data source that exists in the enterprise that you want to be available across systems.

For instance, you might want to share feeds from your documents repository or Factiva news source. The closest analogy of this service on the web is Feedster, an RSS search engine and aggregator.

These feeds are used by RSS readers -- basically to help bloggers with their own knowledge creation via weblogs.

## Basic concepts before we begin (continued)

- **Enterprise aggregation services**
  - Archives collected weblog entries for long-term storage and retrieval
  - Provides a richer set of (automatically entered?) metadata for each entered record
  - Provides some semi-automated means of classification via a controlled vocabulary

7

Finally I talk about **enterprise aggregation services**.

This is typical systems work of aggregating and redistributing data from various sources, but now applied to weblog output. Again, this like Feedster.

## Q: Why weblogs for KM?



8

I think it's important to note **some of the reasons people are using weblogs for KM** in the first place and why **this can be a good thing** for your organization.

While the ease of installing the technology to publish a weblog has a lot to do with the increase of personal web sites, the drive to self-publish probably has more to do with the need to share and incubate ideas, to grow knowledge within communities, and to make connections.

These needs aren't new, but there are reasons why the current trend towards weblogs are a good idea.

So... why weblogs for KM?

## A: Diversity is good.

- **Nardi & O'Day on information ecology**

- A system of people, practices, values, and technologies at work in a local environment.
- A healthy ecology is one that is dynamic, diverse. Diverse set of people and technologies work in a complementary way.

- **ClueTrain Manifesto**

- Bottom up **conversations** are a good thing.



9

### Reason #1 is that diversity is a good thing.

You may know **Bonnie Nardi and Vicki O'Day's article** on Information ecology. They introduced the idea that the business landscape could be viewed using the analogy of ecologies.

The thinking here is that in an information ecology, a diverse number of entities can exist. A diverse set of people and technologies working in a complementary way makes for a healthy ecology.

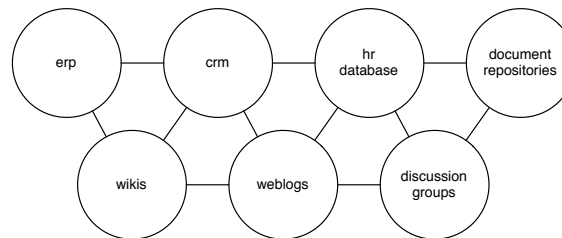
If you're familiar with the **ClueTrain Manifesto**, those authors also discuss locally controlled knowledge as a good thing.

They say the businesses that will succeed are those that enable individuals to converse directly with the organization and with customers.

**The key message here** is that bottom up conversations are good thing.

## A: The new KM is a bottom up effort

- Deloitte: Bridging the gaps between people and systems depends on first creating the conditions that **allow people to participate in KM locally** rather than enforcing technology-based KM policies. These local activities are bridged in loosely coupled Knowledge Networks.
- Forrester: Organizations have begun **to move away from single-solution KM** packages.



10

**Reason #2, Because KM is a bottom-up effort, it's probably more likely to succeed than top-down propositions.**

And market researchers had this to say about KM and software.

**Deloitte** talks here about using knowledge systems to bridge locally created knowledge. This is the idea of loosely coupled systems.

**Forrester** is noting that organizations have begun to move away from single solution. top down KM tools.

### What's the view from the top?

- The natural progression in an information ecology where k-loggers start to proliferate is to seek **a system that pulls together the disparate k-log data.**
- The **role of the information services organization** is to glue together the aggregate of produced k-log data for its users to consume.
- The standards that connect k-log data (XML) will make services more easily achievable.

11

**So with these ideas in mind, what does this mean to libraries?**

It means that we're **well-positioned** to help our organizations connect the pieces in our information ecologies **by taking on the role of aggregator and disseminator** of this locally produced knowledge.

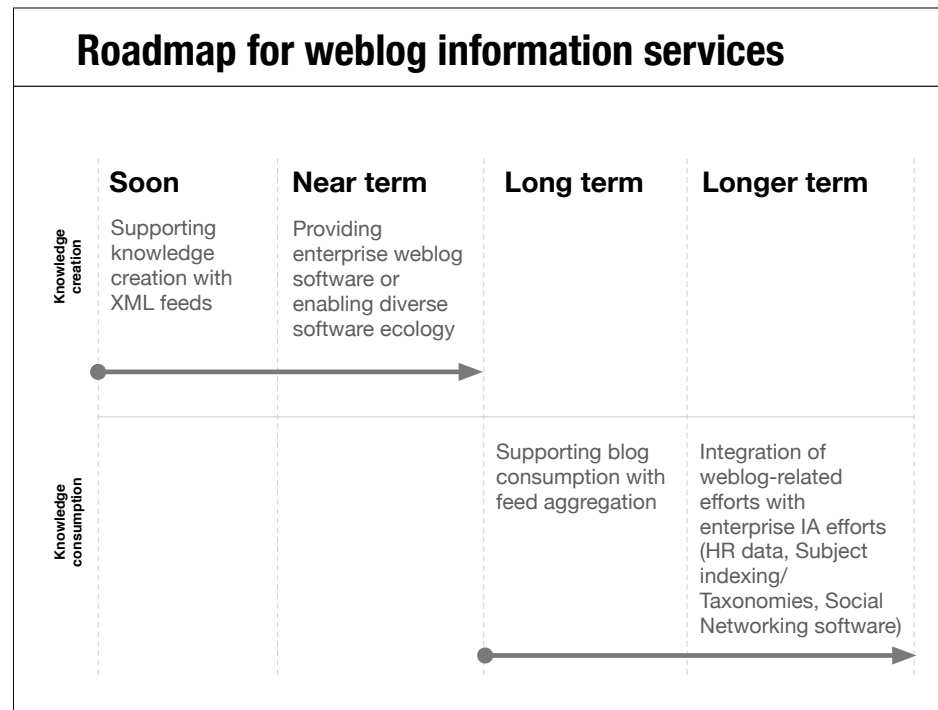
- \* pulling together means merely grabbing data using XML
- \* part of this gluing may include organizing or indexing the data
- \* and finally we'll need to find ways to integrate this data into our current systems



12

Now we're going to talk about the possible stages for integrating blog output into systems.

We'll also take a look at how weblogs relate to other types of enterprise information.



13

So this is the roadmap. These are milestones -- the large targets. This is REACTIVE strategy -- a way to prepare for the growth of weblogs.

\* **At the near end** we are grappling with how to support knowledge creation -- and that's by providing feeds of our data or tools for publishing.

\* **In the middle** we're attempting to provide basic distribution of the knowledge in the enterprise

\* **At the far end**, we're approaching an infrastructure of loosely joined information systems to relate and connect knowledge to various information management resources such as taxonomies and are looking for ways to perhaps better enable active use of these assets with social software

On the way to the far off points, we have a lot of cultural aspects to consider, which may affect your mileage and alter your course. But this is a basic plan towards that integration at the far end.

## Weblogs, guerrilla content in the enterprise

- Lou Rosenfeld's **Roadmap for Enterprise Information Architecture**:
  - Conventional methods for centralized IA efforts -- taxonomy design, site organization, metadata and content management, search usability
  - **"Guerrilla" IA** -- grass roots content produced by experts without necessitating Corp. IT involvement (e.g. Weblogs and Wikis)
  - If Metadata, CMS and Search are core tools of EIA, **K-Log aggregation is the core tool** of "Guerrilla" EIA.

[http://louisrosenfeld.com/home/bloug\\_archive/000212.html](http://louisrosenfeld.com/home/bloug_archive/000212.html)

14

For me, this roadmap is a component of a larger information architecture vision for the enterprise.

You know the conventional methods for centralized IA. They include taxonomy and thesaurus design, metadata, content mangament.

At its core, weblogs are bottom up efforts. The narrow view presented by my roadmap shows how weblogs might fit into the enterprise web services model.

This also fits into Lou Rosenfeld's roadmap for enterprise information architecture as a component he calls "Guerrilla IA".

The comparison here is that if things such as metadata and CMS are core tools of EIA, then blog aggregation is the core tool of Guerilla IA.

## Feeding the Guerrillas

- Guerrilla IA: **keep knowledge creation close to the owners/creators of that knowledge**
- Result of widespread Guerrilla IA: **diverse "information ecology"**, decentralized content sphere with disparate users and technologies behind the content
- Diversity supported by **knowledge networks** created out of EIA efforts

15

The advantage of Guerilla IA is that it keeps knowledge creation close to the knowledge creators. This makes blogs **as intimate as one to one conversations**. This closeness of the owner to the product -- the output -- is what might help sustain weblogging as a model for individual KM.

The result of this form of knowledge creation is a diverse information ecology with data creators **all over the enterprise**.

We're working with the spread out model here and bringing things back together.

## The techno-social life of weblogs

- **Technologically simple; Socially complex**  
(Lilia Efimova)
- Consumer weblog applications are simple / enterprise weblog apps need not be.
- Complexity allows for more granular classification, better commenting in context.
- Communities / social connections may not be as common with enterprise intranet blogs as with public Internet blogs.

16

Here's an good point made by Lilia Effimova, a PhD student who's profilfic on the topic of blogs.

Grassroots creation can be amazingly simple as regards technology.

While the tools themselves are simple, they also offer a degree of complexity with regards to classification and creating social connections that makes them powerful applications in the hands of even novice users.

This is part of the **reason why they're used so much** and part of the reason why **we want to exploit** them as vehicles for knowledge management and business networking.

**So let's look at the types of connections** that are possible with these tools.

## The interconnectedness of blogs

- At some level, most blogs encourage and facilitate interconnectedness
- Simple connections
- **Comments** - User thoughts and annotations attached to blog entries
- **Blogrolls** - A blogger's lists of related blogs in their reading circle

### Comment example in a blog

1. Click here to read or add your own comments

Dec 02, 2003: Skip This Raht and Read Shirky

As you can tell from my title, I'm a fan of this short but brilliant piece Shirky: "The Semantic Web, Syllogism, and Worldview". Shirky shreds assumptions behind the Semantic Web by exposing a major crack in its foundation: a reliance on syllogisms ("If A=B, and B=C, then A=C"). Packages of logical goodness may sound nice on paper, but often require huge generalizations and don't occur frequently in real life, where the S is supposed to actually help us. [More](#)

[permalink](#) | [email this entry](#) | [5 comments](#)

### 2. Comments viewed in context

Comment: [Adrian Howard](#) (Dec 2, 2003)

I'd take a look at <http://www.poorbuthappy.com/ease/semantic/> - a along with the many, many responses.

I'm with the "mostly a straw man" camp. I've yet to meet anybody w web" matches Shirky's definition.

### Blogroll example

| Outbound links: Practice: | Outbound links: Theory: |
|---------------------------|-------------------------|
| > anti-mega               | > joel on software      |
| > laslash                 | > brushstroke.tv        |
| > tomalak's realm         | > in-flight correction  |
| > useit.com               | > glish                 |
| > uidesign.net            | > a list apart          |
| > jlg.net                 | > anne galloway         |
| > peterme.com             | > v-2.org               |
| > elegant hack            | > tesquiden             |

Here's some screenshots showing how people build connections using built in tools such as comments or by manually listing related subject area experts in blogrolls.

The first example here shows how comments are added to a blog. This is Victor Lombardi's IA blog.

The second example shows how an author points to their favorite blogs. This is Matt Jones' IA blog.

## The interconnectedness of blogs

- More complex connections
  - **Distributed discussion**  
(Trackback, links to referrers, embedded search engine query links, **waypath.com**)
  - **Popularity indexes** (MIT/ Blogdex-type aggregation)
  - **Topic-based indexes** (DiceLaRed)

### Waypath links for blog entry

ICANN! NO, YOU CANNOT...

As reported in an article in The New York Times, [Paul Twomey](#), was excluded from a preparatory [Society](#).

Jennifer L. Schenker, of the International Her to come under heavy fire at the conference, v gatherings of high-level government officials, future. An important point of debate will be w American groups like Icann."

According to this article, the role of governme this conference. Although Icann and the US G [Negroponte](#), [Esther Dyson](#) and [Tim Berners-L](#) based companies, including [AOL](#), [Microsoft](#), [Th](#)

K-Collector Topics: [AOL](#) [Microsoft](#) [New York Times](#) [IC](#) [google.it](#) | [permalink](#) | [comments](#) | [trackback](#)

RELATED POSTS FROM [WAYPATH](#)

[Furdlog :: International Technology Policy \[Furdlog\]](#) - Friday, December 12, 2003

[UN taking over the Internet \[Synthstuff+-+music%2C+photography+and+more..](#)

Many may not know this but the United Nations has Currently that power is licensed by the U.S. governm [\[CHUBOGG'S HOUSE!\]](#) - Friday, December 12, 20

18

More complex connections can be seen by embedding related content on other sites using plug-in tools such as waypath, by allowing others' to ping your site using another method called trackback, or by using the Google API to fetch URLs by keyword or using blog entry titles.

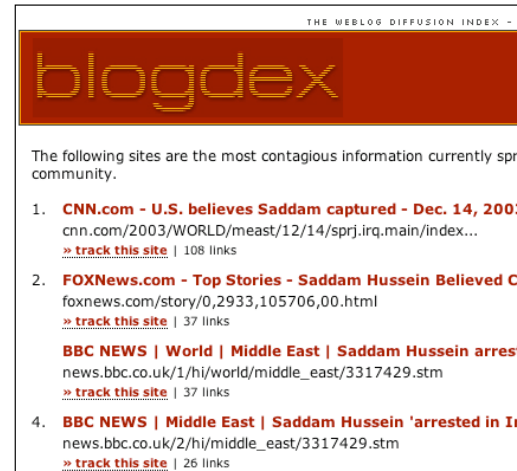
The example here shows how links are embedded into a blog entry by Judith Meskill using Waypath..

## The interconnectedness of blogs

**DiceLaRed** -- dashboard showing most popular topics in a given aggregate of blog entries



**Blogdex** -- index of most popular URLs referred to in a given aggregate of blog entries



19

Topic-based indexes such as DiceLaRed can watch a stream of blog entries and show which topics are most popular within a watched set on any given day.

Popularity indexes such as MIT's Blogdex are used to follow the most linked to blog entries on any day.

## Integrating blogs with enterprise IA efforts

- Connecting people...
  - **to build** Communities of Practice
  - **to leverage** other people's expertise around topics, projects or customers
- Start by understanding information holders and their relationships to information...
  - **by watching** what they publish and how they use information

20

If we take a step back, we start to realize that **making connections** is really a big part of what weblogging is about.

From a systems perspective, there are ways we can help people make further connections by **watching for the explicit relationships found in information we already have** -- by watching what people publish and how they use information.

From there, we might be able to connect authors to topics, users to business units or projects, library users to topic tracking interests.

| Relationships extracted from blog output        |                                                                                                                                                                                                   |                                                                                                    |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| Extraction type                                 | Relationships                                                                                                                                                                                     | Supporting technologies                                                                            |
| <b>Explicit</b><br>Author-created relationships | <ul style="list-style-type: none"> <li>• Blogrolls</li> <li>• Friends on social networking sites</li> </ul>                                                                                       | <ul style="list-style-type: none"> <li>• FOAF</li> <li>• XFN</li> <li>• Social software</li> </ul> |
| <b>Emergent</b><br>Observable relationships     | <ul style="list-style-type: none"> <li>• Comments</li> <li>• Trackbacks</li> <li>• Link logs</li> <li>• Links in blog entries</li> <li>• Reverse-lookups</li> <li>• Popularity indexes</li> </ul> | <ul style="list-style-type: none"> <li>• Social network analysis</li> </ul>                        |

21

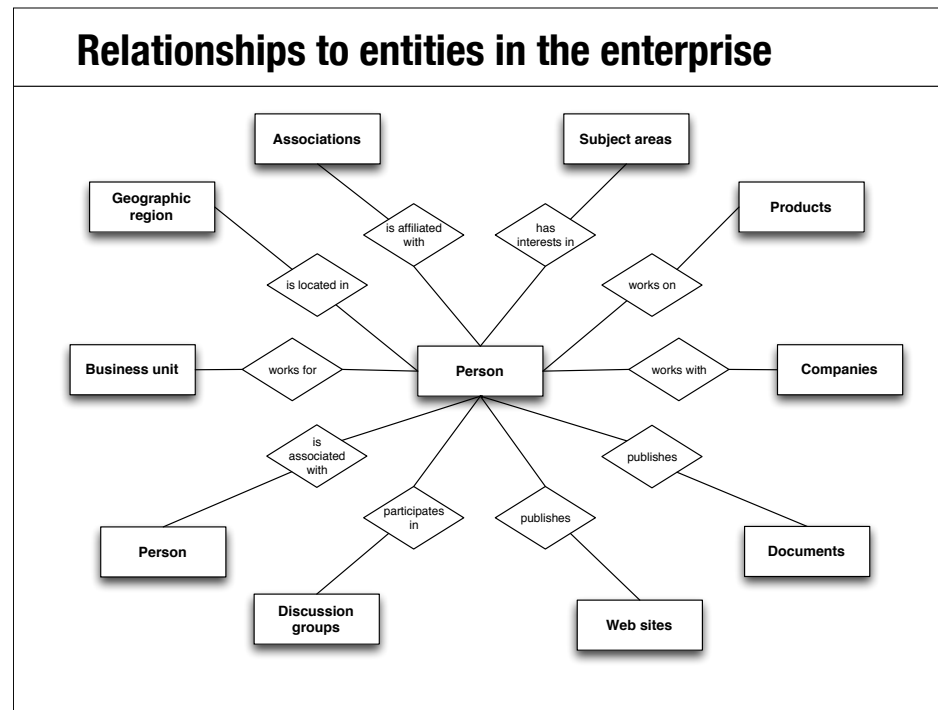
This chart shows the types of relationship that might be extracted through blogs.

**The left column** shows the extraction type or method.

- Explicit relationships identified by authors. This is a knowable realm.
- Emergent relationships are those that can be observed through patterns of pointing and referring, done by the author and by readers. This is a discoverable realm that relies heavily on machine parsing and analysis.

The **middle column** shows where relationships are being extracted from.

The **right column** shows some technologies supporting this extraction.



22

If we broaden our view, we might consider other relationships that can be found by watching the interests of users in information nuggets they publish, in topic tracking interests they record, etc.

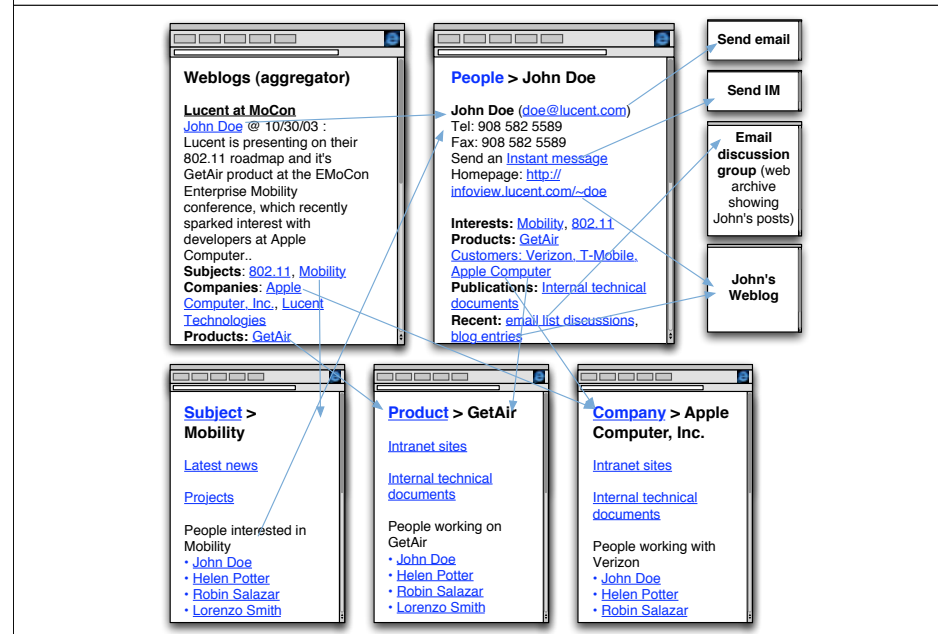
We can turn patterns of use into describable entities.

The idea is to clearly define those entities and create rules for the extraction of those entities from various forms of published matter.

For instance, take a hypothetical person named “**John Doe**”. Let’s say John is interested in “WiFi as a topic” We might know that because he’s published papers on the topic in one of our known sources.

What we need to do is let the system look for these cues and to distinguish that “**John Doe**” is a known person and that “**WiFi**” is a known topic and **connect those 2 entities**.

## Connecting entities



23

I think **this is what makes weblogs interesting** to the organization -- **integrating different types of knowledge** in the enterprise might bring **communities and individuals closer together** around shared interests.

This screenshot represents how this far off goal might look to users.

We're creating knowledge networks that are loosely coupled by allowing systems to talk to each other.

We're **extracting** bits of explicit knowledge and **inferring and suggesting the relationships** of people to concepts, concepts to concepts, and people to people.

This could be the role of the knowledge-based information system.

## Getting there: Which way do you go?



24

Those previous slides should tell you that your roadmap must be able to bend with the changing weblog environment. There are clearly social and cultural elements that will affect how you react to a grassroots knowledge ecology.

The roadmap is pretty simple, but the process and resources needed to get there might not seem so trivial. So let's look at some approaches.

### 3 approaches on the weblog services roadmap

- Seems to me there are **three different approaches** or strategies
  - Centralized
  - Decentralized
  - Mediated

25

We're going to look at three different approaches, starting with a centralized approach.

### **Centralized approach**

- One enterprise weblogging and news feed aggregation application
  - Integrated news aggregator
  - Integrated classification / controlled vocabularies (subject, author. etc.)
  - Integrated publishing

26

This approach is one where publishing, organizing, and distribution are centrally controlled.

### **Success factors for centralized approach**

- Demands standard processes / ways of working
  - Likely to work in smaller organizations
  - In large organizations, understanding processes of all user types and communicating the value of centralization is difficult
- Requires a lot of user research, continuous communication and training

27

Here are some factors needed to succeed with this approach.

The centralized approach requires standard processes, so this means understanding how people work.

This probably means this approach works best with smaller organizations.

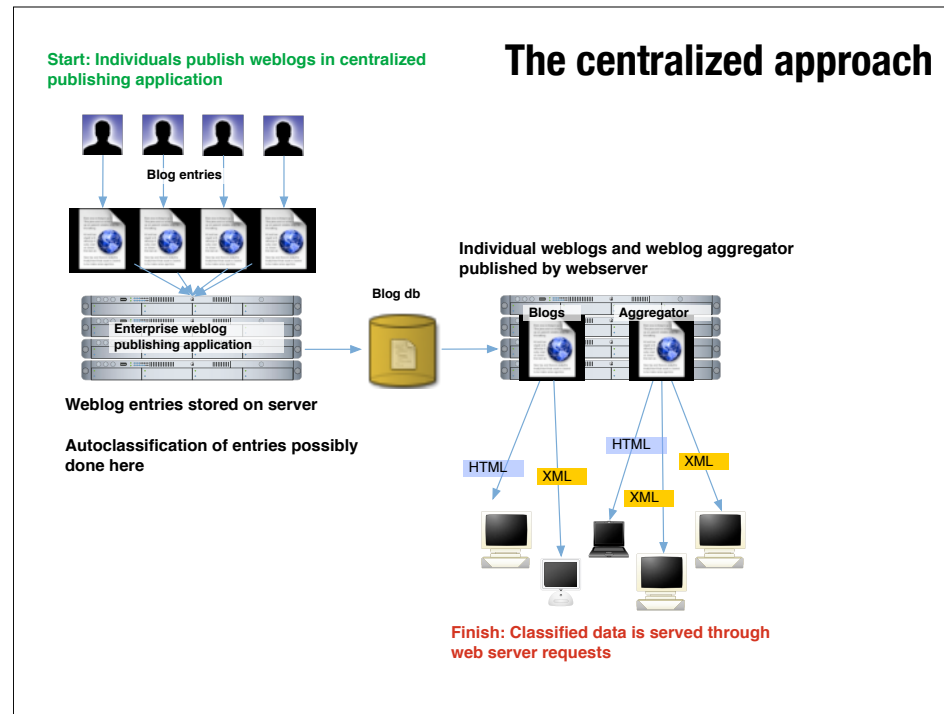
This approach also requires a lot of user research and continuous communication and training.

### **Benefits for centralized approach**

- May provide ability to use more complex application features
  - Email based blogging, contextual commenting, contextual tagging
- High returns in ease of administration

28

A centralized approach makes administration easier and allows more complexity in terms of features. I'll show some examples of this in minute.



29

Blog content creation goes in through a centralized tool or website.

The data gets parsed and autoclassified and then served. One advantage of a centralized tool is that bloggers could use a single controlled vocabulary for adding index terms to their entries.

Creation, publication and distribution are all done through one system.

# Centralized example: Enterprise blogging applications

---


Front Page > InBo

July 1, 2002 - October 9, 2003
 Full ▾

**Active Years**

2002

2003

Search

---

advanced search

 Add New Article

 Show Collector

 Print Version

 Reclassify

 Email Articles

 Collect Articles

 Add Comment

 Email Reply

 Visibility

---

**Shortcuts**

Traction Skins:

- Modern skin
- News skin
- Basic skin (SDL)

Other SDL Skins (to Reset,

**DLL Install Errors on TIP V1** ✉

InBox116: 17 Jun 2003 11:17 am

Posted by Visitor

[Cross References](#)

---

TIP V1 generally installs just fine, but when I install it for our users, I occasionally get a harmless DLL error like "Cannot access c:\winnt\system32\iphalpi.dll" 01 [ProductAlpha: Bug](#)

I click to ignore and after installation the application seems to work fine. Please confirm this is OK. 02 [StarCustomer: do](#)

Thanks, 03 ✉

Joe 04 ✉

Reply-To: frank@tractionsoftware.com  
From: joe.blogs@starcustomer.com  
To: tractioninbox@cox.net  
Date: Tue, 17 Jun 2003 11:16:18 -0400

---

 [Add a Comment on this Article](#)

**Related Entries:**

[referenced by \(1\)](#)

[ProductAlpha74: Product Alpha V2 Requirements](#)

[is replied to by \(1\)](#)

[StarCustomer111: re: DLL Install Errors on TIP V1](#)

## Email to blog

Traction software allows users to blog from email. This example shows a customer service blog that can be used via email or via the web interface.

### Email to blog

Traction software allows users to blog from email. This example shows a customer service blog that can be used via email or via the web interface.

Traction allows people to use email to post to the enterprise blog, and because email is a technology that most people are comfortable with, it's not very disruptive. It doesn't have a steep learning curve.

## Centralized example: Enterprise blogging applications

**IE, Outlook, and Word Application Integration** Feature Integration  
ProductAlpha52: 9 Jun 2003 8:35 am  
Posted by JFrank  
[Cross References](#)

While allowing for basic copy, paste, and move operations from any windows desktop application, TTP has hooks to enable easier publishing from and source references from selected office applications including IE, Outlook, and Word. 01

**Base Requirements:** One common feature to all is ability to select text and graphics, click the TTP icon, see an added section stating "Publish [from Chosen Application] With or Without Original Attached." Choosing attach original will wrap up a copy of the original page or document and attach it to the new article in Traction. 02

Integration Requirements for each are detailed below: 03

**Internet Explorer** 04

1. Base requirement Publishing with Original attaches a single MHT file. 05
2. Source reference including domain and full URL 06
3. Right click publishing option from IE menu 07
4. HTML simplification and cleanup operations 08

**Outlook Integration** 09

1. Base requirement Publishing with Original attaches a single EML file. 10
2. Source reference including email header items: To, From, Date at trail and of the new Traction article 11
3. Attachments are attached 12
4. Smart Tag cleanup operations 13

**Annotating at the paragraph level.**  
This example is using the Traction blog application for collaborative authoring of a product requirements document. Note the status flags

31

Here's another example from Traction showing weblogging can be **a bit more complex**. Their tool allows **annotation at the paragraph level**.

This example is actually using Traction for **collaborative authoring** of a product requirements document. Note the status flags. users here are actually changing the status of bullet point as they evolve this document.

When you're dealing with one centrally-controlled platform, the advantages in terms of functionality become pretty clear, I think.

## **Decentralized**

- Many diverse / disparate publishing applications including weblogs
- Common XML output format
- XML feed aggregator
- Search service with possible auto-classifier

32

Now the decentralized approach. This approach allows many blogs to exist anywhere.

It relies on common XML outputs for collection.

Brings that together with a feed aggregator.

And provides access via a search service with possible auto-classifier.

### **Success factors and benefits**

- Perhaps the easiest to execute
- Supports diversity; gives users freedom to choose tools that match their processes
- Requires the least amount of effort and human resources on the system side
- Requires the considerable financial investment
- May not provide the most relevant results through machine classification

33

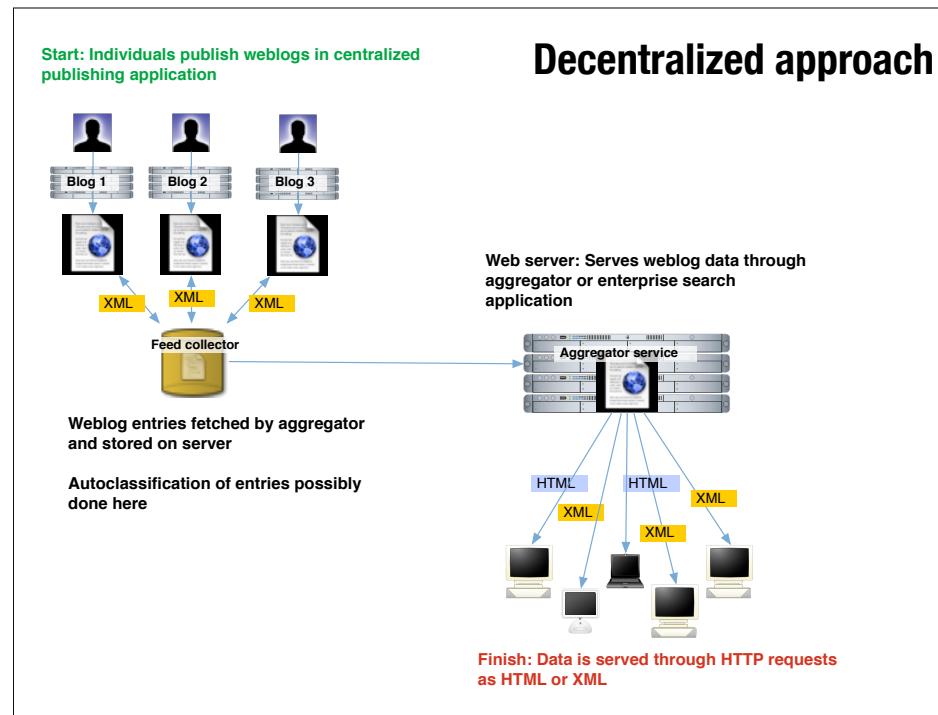
This is perhaps the easiest approach to execute.

It supports diversity and let's users choose their tools.

It requires the least amount of effort and human resources.

But it requires a good deal of initial financial investment

and may not provide the most relevant results because we're relying on machine classification. The tools you can buy for the search and retrieval are getting better all the time, however.



34

In the decentralized approach, knowledge creation happens anywhere and a machine collects the XML output.

Some autoclassification might be done when the feeds enter the system.

Then the server redistributes the aggregate when people search.

## Centralized example: Corporate data as RSS

Home | MyView | **Collections** | Topics | Databases | About InfoView | IIS services | Sitemap | Contact us

Home > Collections > External news services > Factiva News and Business Information Service

### Search results for Factiva News and Business Information Service

[E-MAIL RESULTS](#) [BROADEN SEARCH](#) [XML](#) [FULL](#) [RSS BRIEF](#)

Search Terms: [knowledge; management](#) -- 14215 records retrieved  
TopicTracker: [Set up a daily alert saving this search strategy.](#)

Add terms to reduce the number of retrieved records:

[New Search](#) | [Search Help](#)

If you did not find what you wanted, call our helpline at 908 582-4840 or email [solution@lucent.com](mailto:solution@lucent.com) for further assistance.

**Narrow your search using a major topic category present in the retrieved documents**

- [Knowledge Management](#)
- [Customer Relationship Management](#)

**Sort by:**

- Establishing a Knowledge Management Culture Expands Consumer Base and Spurs Growth; *Business Wire* November 12, 2003
- Prophet 21 Releases Knowledge Management Center 3.5 for CommerceCenter; *Product News Network* November 4, 2003
- ClearForest Named a Finalist for 'KMWorld Promise' Award Content Extraction and Analysis Provider One of 10 Companies Recognized by KMWorld Editors for Delivering Truly Innovative Knowledge Management Solutions; *PR Newswire* November 12, 2003
- NGRAIN-TM- Software Brings Advanced 3D Capabilities to Product Knowledge Management; *Business Wire* November 3, 2003
- KNOWLEDGE MANAGEMENT MART SET TO BOOM; FROST and SULLIVAN (to touch

#### Fetching database results as RSS

Users can retrieve results from this search strategy periodically using an RSS reader

35

One way we help support a diverse set of bloggers is by giving them RSS feeds for our databases. They use these in their RSS readers to track topics without leaving the firewall.

## **Mediated**

- Many diverse / disparate weblogging applications
- Common XML output format
- Possible shared controlled vocabularies for classification upon blog creation
- XML feed aggregator
  - Some auto-classification
  - Human mediation
- Enterprise search with some auto-classification

36

The last approach involves more human involvement on the systems end.

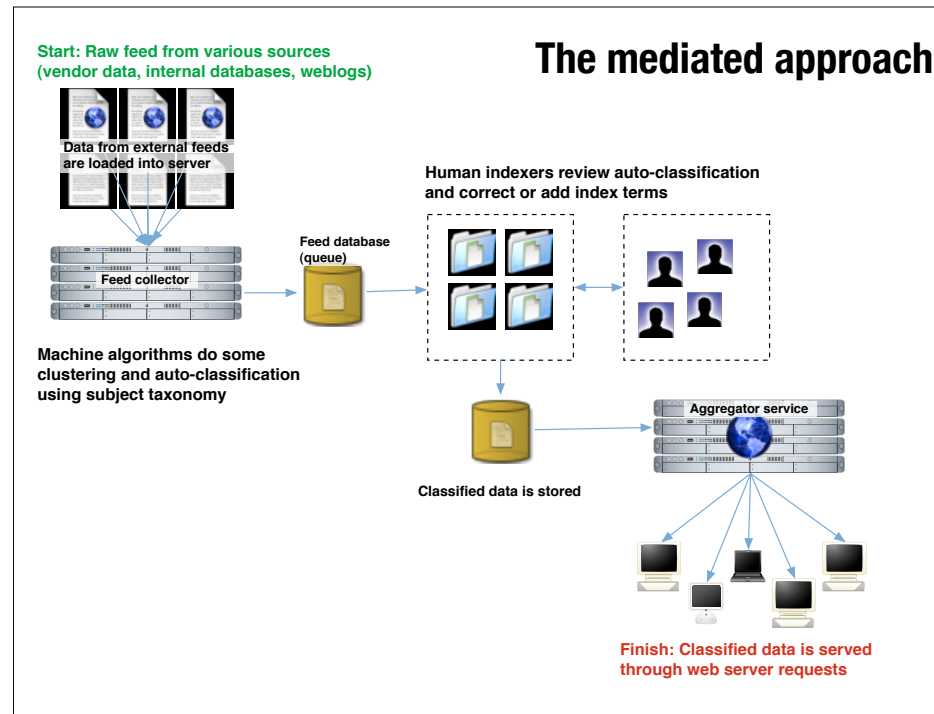
### **Success factors for mediated approach**

- Requires considerable human resources and custom programming
- Supports diversity; gives users freedom to choose tools that match their processes
- Gives users freedom to choose tools that match their processes
- May provide the best (most relevant) results in terms of classification

37

This approach requires more resources, but supports diversity and gives users control of their tools.

This may provide the best results, however, because classification is reviewed and corrected by humans.



38

Using this approach, once again, knowledge creation happens anywhere.

A feed collector grabs XML output and then we do some clustering and auto classification using a subject taxonomy.

The content is reviewed by human indexers and then stored for re-publishing and distribution.

# Decentralized: News aggregator

W4k-collector

Home Today What Where Who About k-collector

## what

(most recently updated of [307 topics](#)):

- [blogging](#) [1314]
- [Weblogs](#) [592]
- [Literature](#) [24]
- [E&T 1.0](#) [30]
- [TV](#) [131]
- [Listings](#) [373]
- [Events](#) [235]
- [Music](#) [208]
- [Design](#) [236]
- [Memes](#) [38]

## where

(most recently updated of [69 topics](#)):

- [Microsoft](#) [440]
- [Apple](#) [147]
- [London](#) [21]
- [Iraq](#) [2]
- [San Francisco](#) [66]
- [Yahoo](#) [53]
- [Dell](#) [25]
- [New York Times](#) [84]
- [BloggerCon](#) [492]
- [Earth](#) [41]

## who

(most recently updated of [206 topics](#)):

- [Paolo Valdemarin](#) [23]
- [Dave Winer](#) [187]
- [Robert Scoble](#) [60]
- [Howard Dean](#) [148]
- [Ray Ozzie](#) [11]
- [Halley Sutt](#) [1]
- [LockerGnome](#) [3]
- [Neo](#) [2]
- [Doug Engelbart](#) [2]
- [Jon Udell](#) [52]

## when

Time and dates:

| November 2003 |     |     |     |          |     |     |
|---------------|-----|-----|-----|----------|-----|-----|
| Mon           | Tue | Wed | Thu | Fri      | Sat | Sun |
|               |     |     |     |          | 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  |
| October       |     |     |     | December |     |     |

## today

AAA

**Curiouser and curiouser!**  
[Where is all the music?](#)  
 Something that has surprised me travelling in and out of London twice a week is how few people listen to music. I mean... [\[continue\]](#)

**Family Medicine Notes**  
[John Faughnan](#)  
 Hey - John's got a weblog -- er -- three. [\[continue\]](#)

**Family Medicine Notes**  
[Teflon Flu](#)  
 Here's an interesting piece on the Teflon Flu. [\[continue\]](#)

**Marc's Voice**  
[Categories to Topics](#)  
 Results from Paolo's first attempts at gluing into Dave Winer's new categories stuff. [\[continue\]](#)

**Sandhill Trek: A Public Space for Self Expression**

**I Knew Things Had Changed**  
 I Knew Things Had Changed... [\[continue\]](#)

**house of warwick**  
[Bryan hits another out of the park RE: the cactus](#)  
 RIP Scripting News Cactus. [\[continue\]](#)

**Blog for America**  
[In The News: Cheney's War?](#)  
 As the war in Iraq continues with no end in sight.

Here's an example of an RSS aggregator called k-collector that pulls weblog output and re-publishes content using a faceted approach.

### Reactive versus proactive strategy

- **Best bet is a reactive strategy (decentralized and mediated):** Focus on watching information publishing and usage behaviors (including weblog-related activity).
- **More burdensome is the proactive strategy (centralized):** Urge people to participate in knowledge creation efforts.

40

Here's my take on coming up with a strategy.

I think my organization would do best to **stay reactive** -- using the **decentralized or mediated approaches**

**I don't think we would ever urge people** to participate in weblogging

As an aside, we're starting to consider the prospect of doing weblogging ourselves to bring alternative sources for information to our customers' attention.

### **Making the most of a reactive strategy**

- **Extract data** in the enterprise; **build knowledge networks** by relating entities
- Capture/mine explicit knowledge stored in: publications, discussion groups, blogs/ personal home pages
- Use the extracted data to connect people to knowledge, people to people (experts, CoPs),

41

I like the idea of capturing information about what our users write in self-published weblogs, in what they write in printed publications, and in the subject preferences they specify in our library system through topic tracking.

I think all of this information can potentially be used to help people do their jobs by connecting them to people and resources.

This way of approaching KM seems diverse and organic to me and may have a good chance of succeeding because of that.

## **Finding solutions**



42

So if you buy all of this, you'll eventually get to the question of "how do we do this?".

### The solution that matches your info. ecology

- **Size** is a factor
  - Small groups may not need large knowledge networks. A centralized approach would work.
  - Large organizations seem to benefit most from decentralized and mediated approaches.
- **Resources**
  - Do you have the financial and human resources to pull off any of these solutions?

43

**Finding the solution that fits your ecology** depends on different factors such as **size**.

It might seem obvious, but **small groups might find it easiest to go with a centralized solution**. If the group is small enough, no system at all might be the best solution.

Large groups seem to benefit from the decentralized and mediated approaches.

Obviously the **human and financial resources** needed to initiate and sustain these efforts needs to be considered.

### **The solution that matches your info. ecology**

- **Culture**
  - Do personal home pages/bloggers exist now?
  - Does your org. have an attitude of openness regarding knowledge sharing?

44

**Culture** plays probably the **largest part in determining how you should proceed.**

We have to **examine what exists now** in the organization and **consider issues about openness** in knowledge sharing.

## The solution that matches your info. ecology

- **Culture**

- Weblog growth with knowledge workers is uncertain (Lilia Effimova)
- Protecting knowledge and self-interests
- Corporate policies on visibility of knowledge assets; competitive intelligence
- When and if it happens... Key to success is easing knowledge workers into the role of knowledge loggers

<http://blog.mathemagenic.com/2003/10/13.html#a796>

45

Some cultural aspects may prevent weblog growth. For instance,

- **a culture of protecting one's own knowledge** and self-interests may prevent people from wanting to blog
- **corporate policies on publishing information** on the intranet may preclude blogging.
- **but when and if the interest in blogging appears**, the key is to ease knowledge workers into the role of knowledge loggers. Constant communication with bloggers is the key to sustaining your relationship with them.

### **Build vs. buy**

- We always **build**
  - Pros: Customization, integration with existing systems, sensible evolution of features
  - Cons: Possible slower development, requires more human resources and training

46

At some point you'll have to decide if you're going to build your own system or buy the technology.

My organization always builds because there are benefits in terms of customization and integration with our existing systems. We tend to evolve our tools slowly over time.

The downside, of course, is that this requires time and considerable human resources.

### **Build vs. buy**

- It's easier to **buy** if you have the resources
  - Pros: Quicker development, possible community of users, technical support
  - Cons: Requires initial investment of money and training, may require investment in consulting services over time

47

And if you have the money, it may be easier to buy.

**Are we there yet?**



### **Where are we now? First steps.**

- Most companies are probably still waiting / evaluating
- First steps
  - Support knowledge creation with XML feeds of databases
  - Use existing enterprise search to make blog data findable in a rudimentary fashion
  - Connect with bloggers (possibly providing a blog bloggers) to list users and evaluate needs

49

I think most companies are probably still waiting and evaluation, but I've been told that some companies are actually buying products like Traction Software.

I think the first step of providing XML feeds with bibliographic data for various enterprise databases is a fairly easy one to start with.

Next you can use your existing enterprise search to make blog data findable in a rudimentary fashion, until you're ready to think of your position in terms of aggregation services.

Connecting with bloggers as they appear is the right way to help keep your strategy based on users' needs.

### **Next steps; dealing with growth**

- As blogs mature, embed them into enterprise IA processes using semi-automated methods for classification
- Continue to investigate user needs and features, e.g. Blog creation services, Social networking needs
- Later, consider data analysis tools and integration of knowledge entities (the far off goals)

### **Vendors to look at**

- For blog creation consider content management vendors focussed on blogs
  - Traction; <http://tractionsoftware.com>
  - Userland Radio; <http://userland.com/>
  - Six Apart; <http://sixapart.com> (Pro Version?)
  - Drupal; <http://drupal.org>
  - .Net developers (because some IT groups are married to MS)

51

Consider scripted/cached models versus static page builders

### **Vendors to look at (continued)**

- For more robust content management, consider CMS vendors with taxonomy focus
  - Stratify; <http://www.stratify.com>
  - Inmagic; <http://www.inmagic.com/>

## **Vendors to look at (continued)**

- If only throwing search at the problem, consider enterprise search vendors
  - Verity K2 Classification Suite; <http://verity.com>
  - Autonomy Enterprise Search; <http://autonomy.com>
  - Endeca ProFind; <http://endeca.com>
  - Convera RetrievalWare; <http://convera.com>
  - InQuira; <http://www.inquiria.com>
  - Applied Semantics; <http://appliedsemantics.com>

# **Thank you. Questions?**

**Michael Angeles**  
**jibbajabba@mailblocks.com**

**Slides available here:**  
**<http://urlgreyhot.com/files/cil-presentation/>**