HPVIS and the "Architecture of Participation"

Presented at
First U.S. Conference on
Characterizing Chemicals In Commerce:
Using Data on High Production Volume Chemicals
December 14, 2006

Scott Butner
Senior Research Scientist
Knowledge Transformation and Integration Group
Pacific Northwest National Laboratory

What is "The Architecture of Participation?"

- Term coined by Tim O'Reilley to describe social computing environments in which users are an integral part of adding value to data
  - Perhaps best exemplified via recommendations, reviews on Amazon.com
  - Wikipedia is another popular example
- "Architecture of Participation" is a fundamental element of what has been dubbed "Web 2.0"
- A philosophy regarding data, users, and application development

Why is this Relevant to HPVIS?

- Brion Cook's goal of "…a database that meets everybody's needs" is laudable
- But it's not *possible*!
- Web 2.0 concepts – rooted in the Architecture of Participation – can move us closer to the realization of that goal than a monolithic application
  - By opening up development routes never anticipated
  - By harnessing the collective energies and efforts of a much broader development team
  - By facilitating integration of HPVIS data with other information

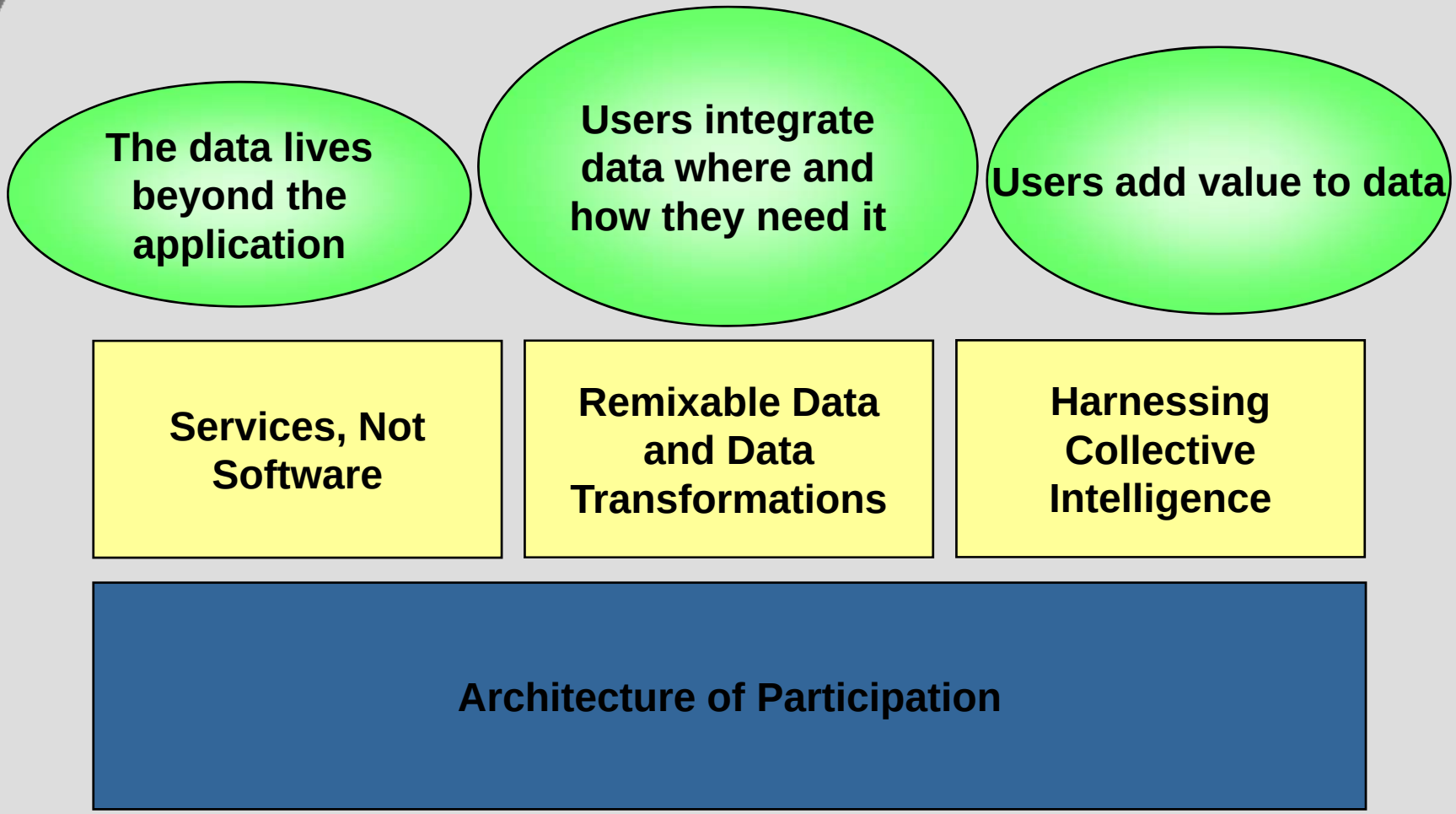
What are the Key Themes Underpinning Web 2.0?

O'Reilly identified the following key themes:

- Cost Effective Scalability
- Software Above the Level of a Single Device
- ***Services, not Packaged Software***
- ***Remixable Data Sources & Data Transformations***
- ***Harnessing Collective Intelligence***
- ***"Architecture of Participation"***

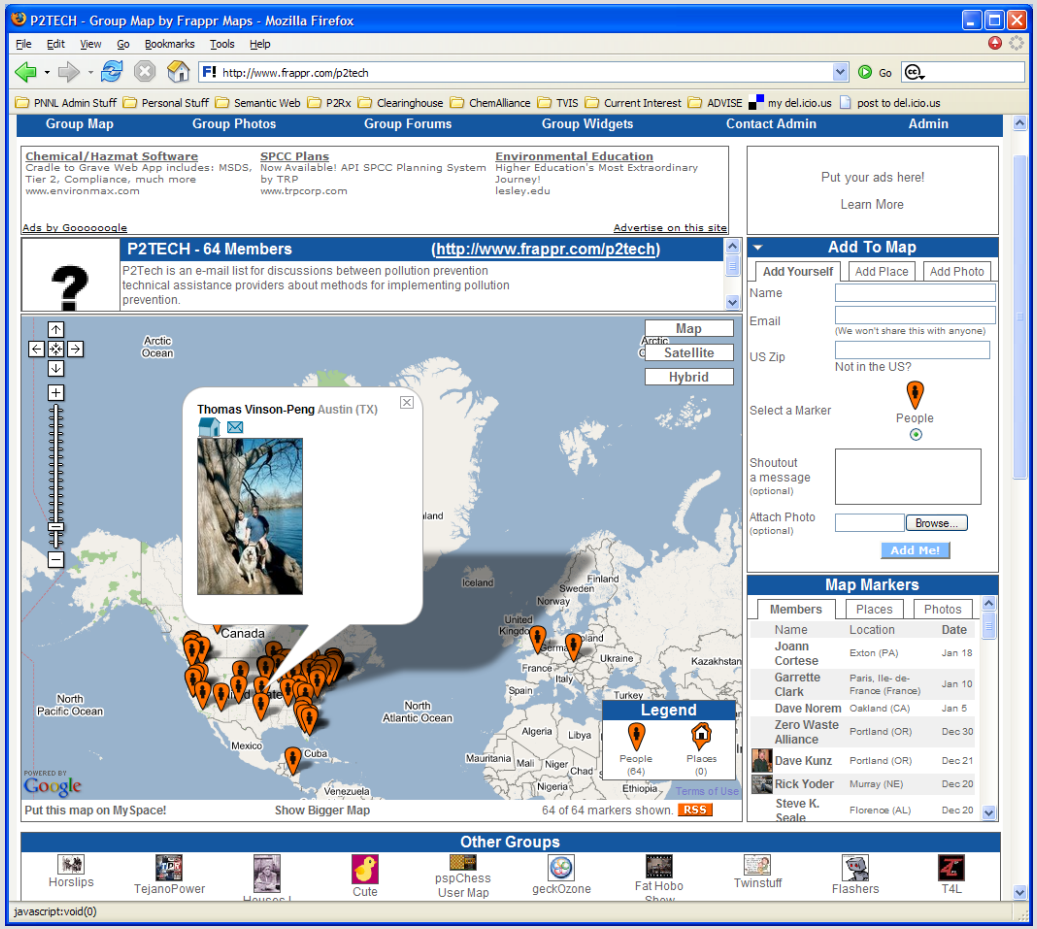
Highlighted themes are probably the most critical to HPVIS

Elements of the Architecture



Services, not Packaged Software – Web-based API's and Mash-ups

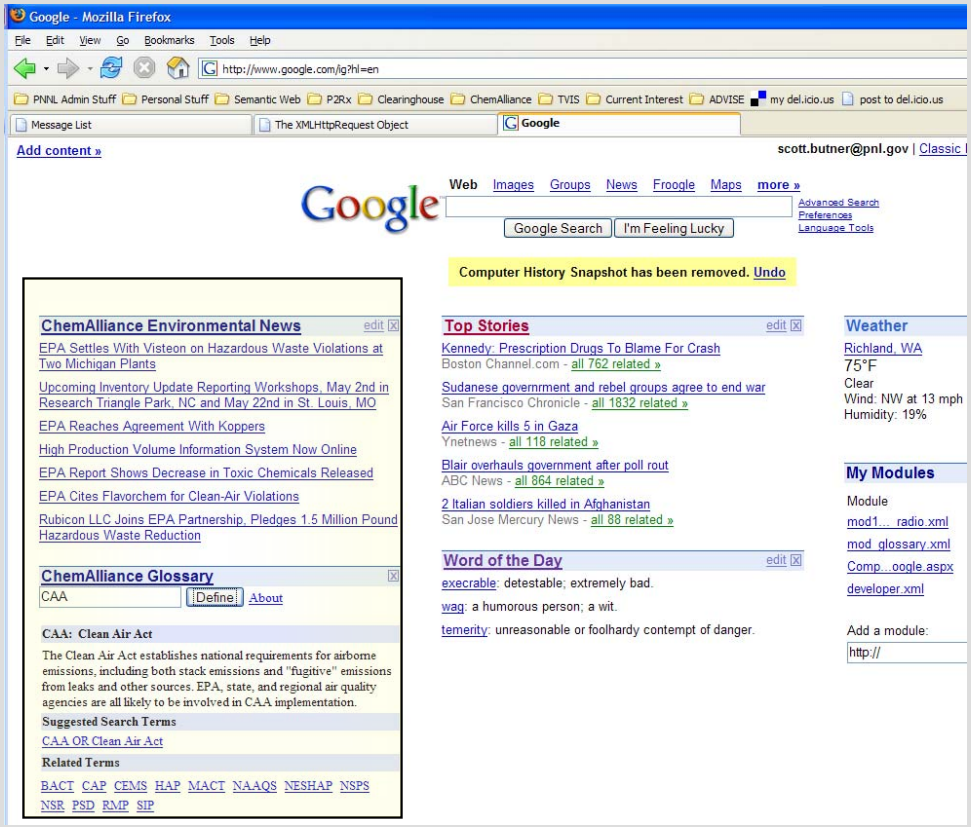
- Web-based APIs (Application Programming Interfaces) provide access to sophisticated capabilities without heavy programming requirements
- Can be combined in “Mash-ups” to extend capabilities or add value to data



Frappr -- http://www.frappr.com/p2tech

Remixable Data Sources & Data Transformations

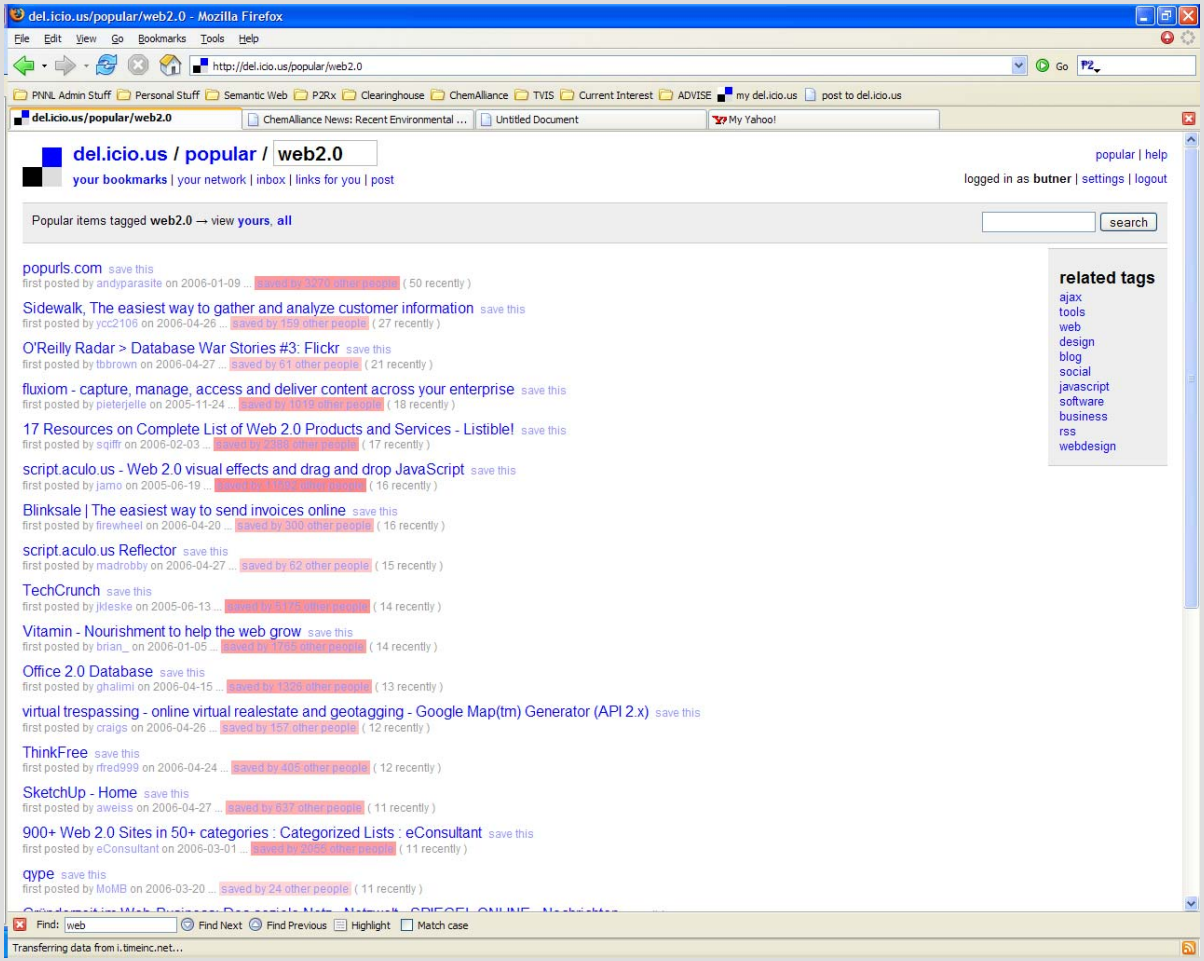
- Modular content and web services allow users to “mix and match” their own content
- Allows content providers to get a place on the daily desktop



Google personalized homepage – easy modular content
http://www.chemalliance.org/webservices/modules/mod_glossary.xml&q=chemalliance&start=0

Harnessing Collective Intelligence – Folksonomy & Tagging

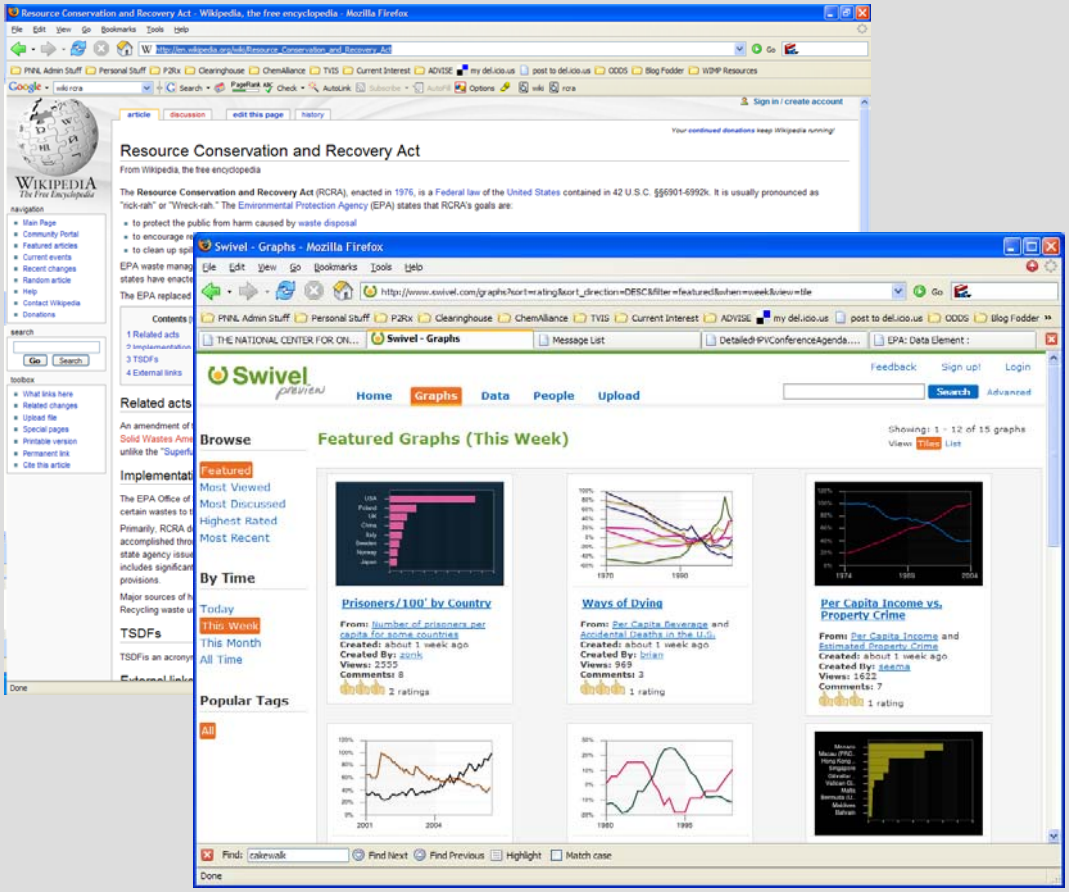
- Turns the idea of "controlled vocabularies" upside down
- Users decide how to label information; convergence on terms ("tags") is an evolutionary process



http://del.icio.us/ -- a prominent social bookmarking site

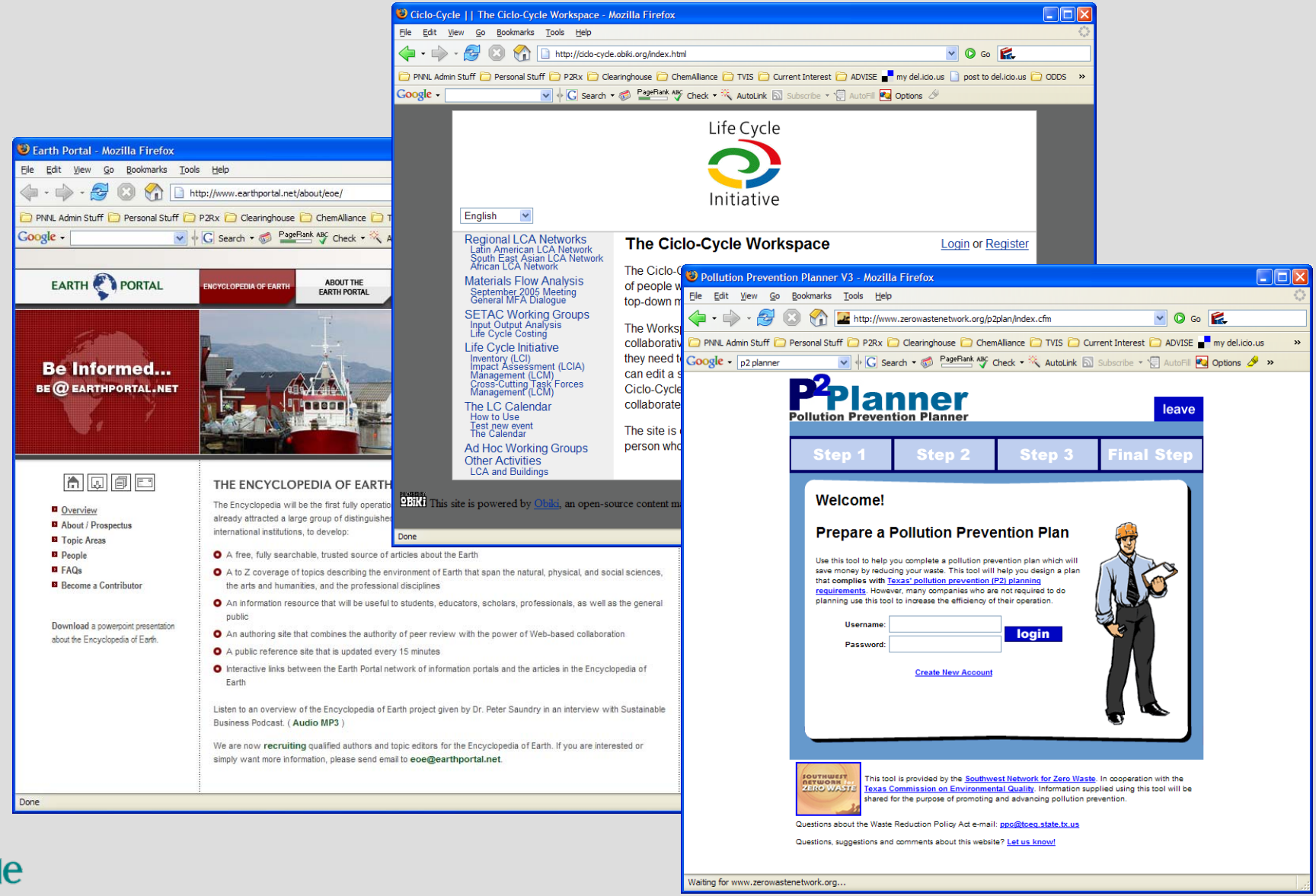
"Architecture of Participation" – Wikis and Blogs

- Web 2.0 puts the ability to create content – for better or worse – into the hands of users
- By simplifying and democratizing web content development, it creates new opportunities and new challenges for "authoritative sources"
- Realize that this is going on with or without you and your program!



Wikipedia entry on RCRA -- http://en.wikipedia.org/wiki/Resource_Conservation_and_Recovery_Act

Early Adopters in the Environmental Domain



Battelle

Applying the "Architecture of Participation" to HPVIS

- Enable service-level access to the data
  - Ad-hoc queries should support web services interface, XML outputs, and some degree of self-documentation
  - Support for queries that enable mash-ups:
    - List of all chemical names, mapped to ID
    - Query by CAS
- Link to folksonomy-based "workspace"
  - Users need the ability to link HPVIS data with other information resources
  - Folksonomy would allow new categories to be defined
    - E.g., by product, location (facility), project
- Support for visualizations of data
  - Compelling applications of the data are often facilitated by compelling presentation of the data

Applying the "Architecture of Participation" to HPVIS (continued)

- Harness collective intelligence and social participation more effectively
  - Much of the data in HPVIS is narrative and/or bibliographic in nature
  - Lends itself to wiki format quite nicely!
- Make the semantics explicit
  - At the very least, publish a data dictionary!
- Harness the spirit of "perpetual beta"
  - E.g., Google's "Summer of Code", Yahoo's API library
  - Open doors to making your user community part of the development team

Conclusions

- Some – maybe the best – applications of HPVIS data have not yet been envisioned
- Emerging trends in web technology facilitate much greater interaction between user community and the data they need
- HPVIS tool as configured does little to support this approach
- “opening up” the data can be done without abandoning the existing investment in data, technology