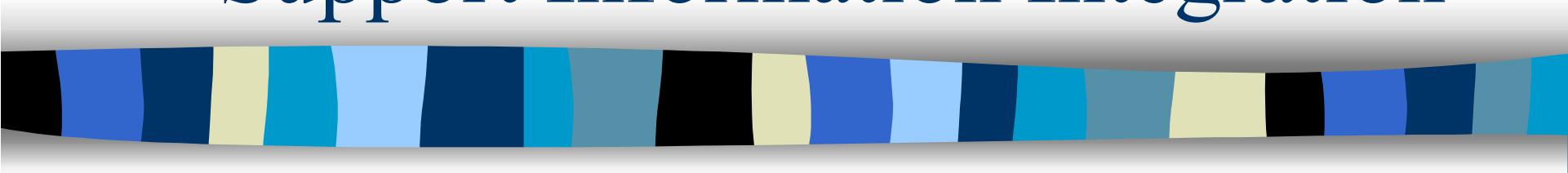
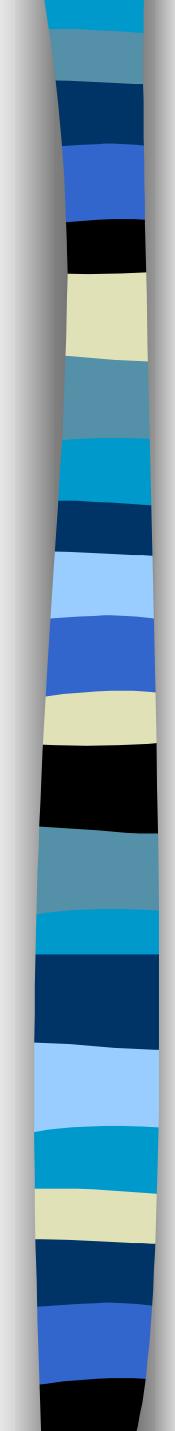


Using Open Hypermedia to Support Information Integration



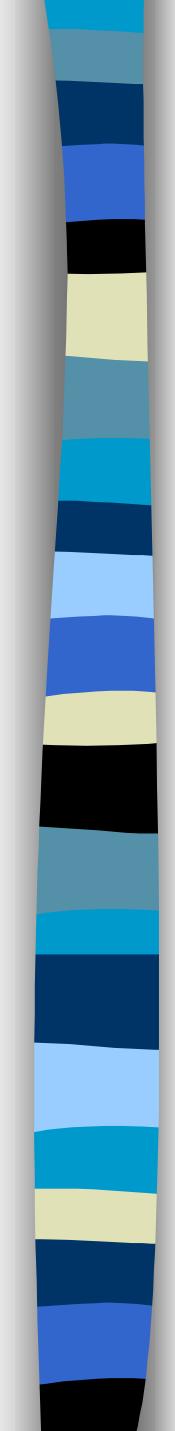
Kenneth M. Anderson and Susanne A.
Sherba

University of Colorado, Boulder
{kena, sherba}@cs.colorado.edu
August 15, 2001



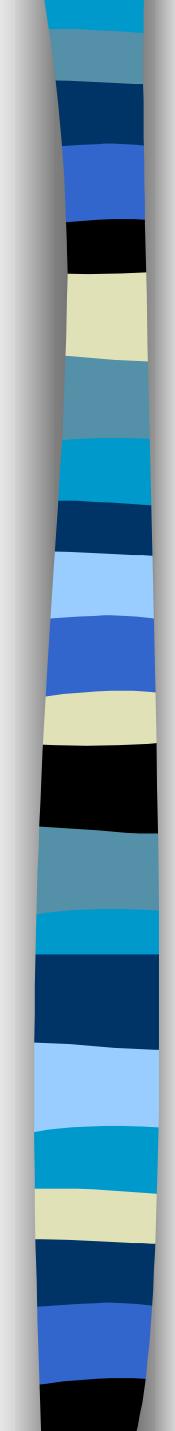
Overview

- Information Integration
 - Managing software artifact relationships
 - InfiniTe Information Integration Environment
- Open Hypermedia
 - Design Influences
 - Delivery mechanism for InfiniTe results
- Demo and (brief) Conclusions



The Problem (still)

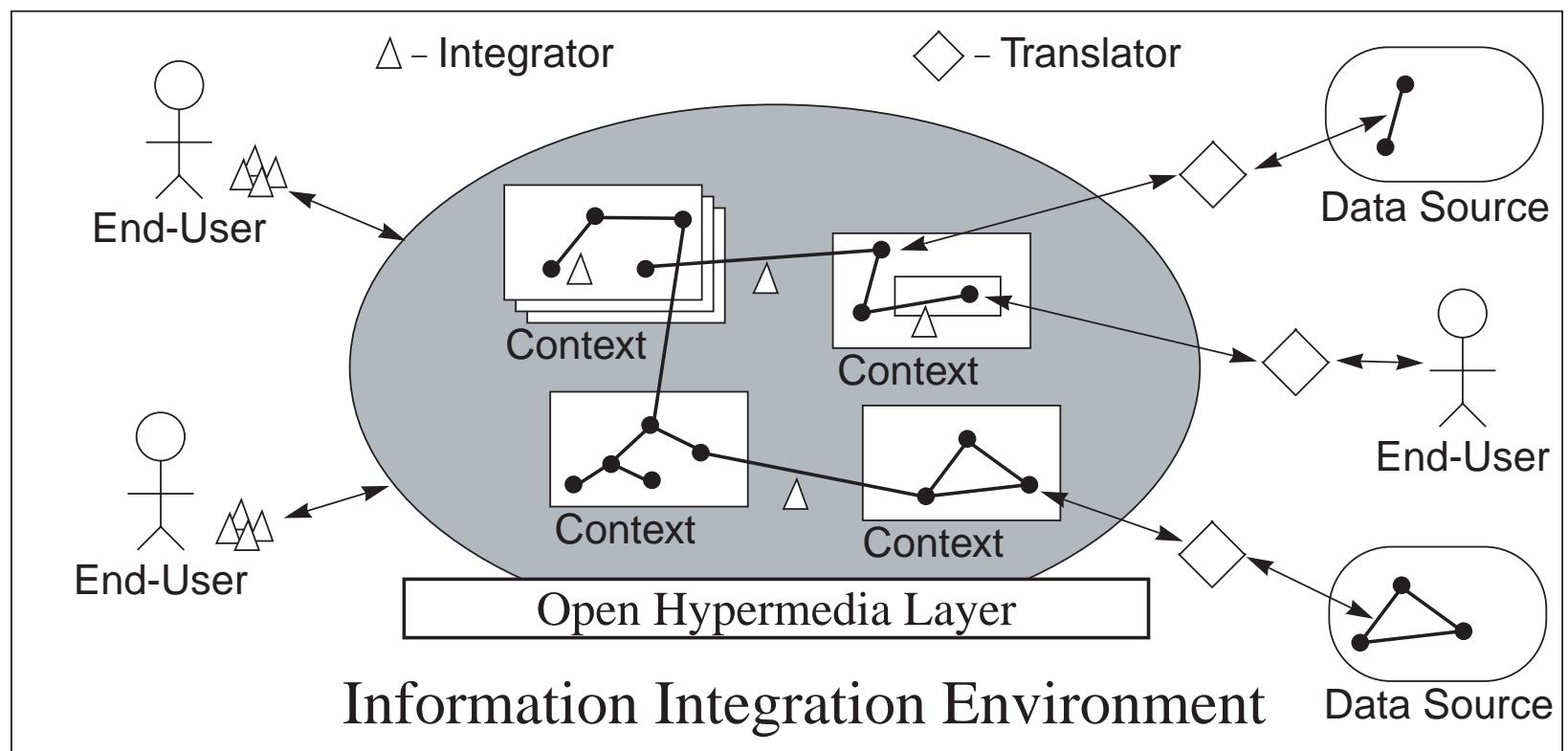
- Software artifacts contain a multitude of implicit and explicit relationships
 - Northrop Grumman's B2 avionics software
 - nearly 600,000 instances of six different link types
 - across 34,000 pages of software artifacts and
 - 100,000's of lines of source code
- Open hypermedia research has been usefully applied to these problems, but has not provided a complete solution



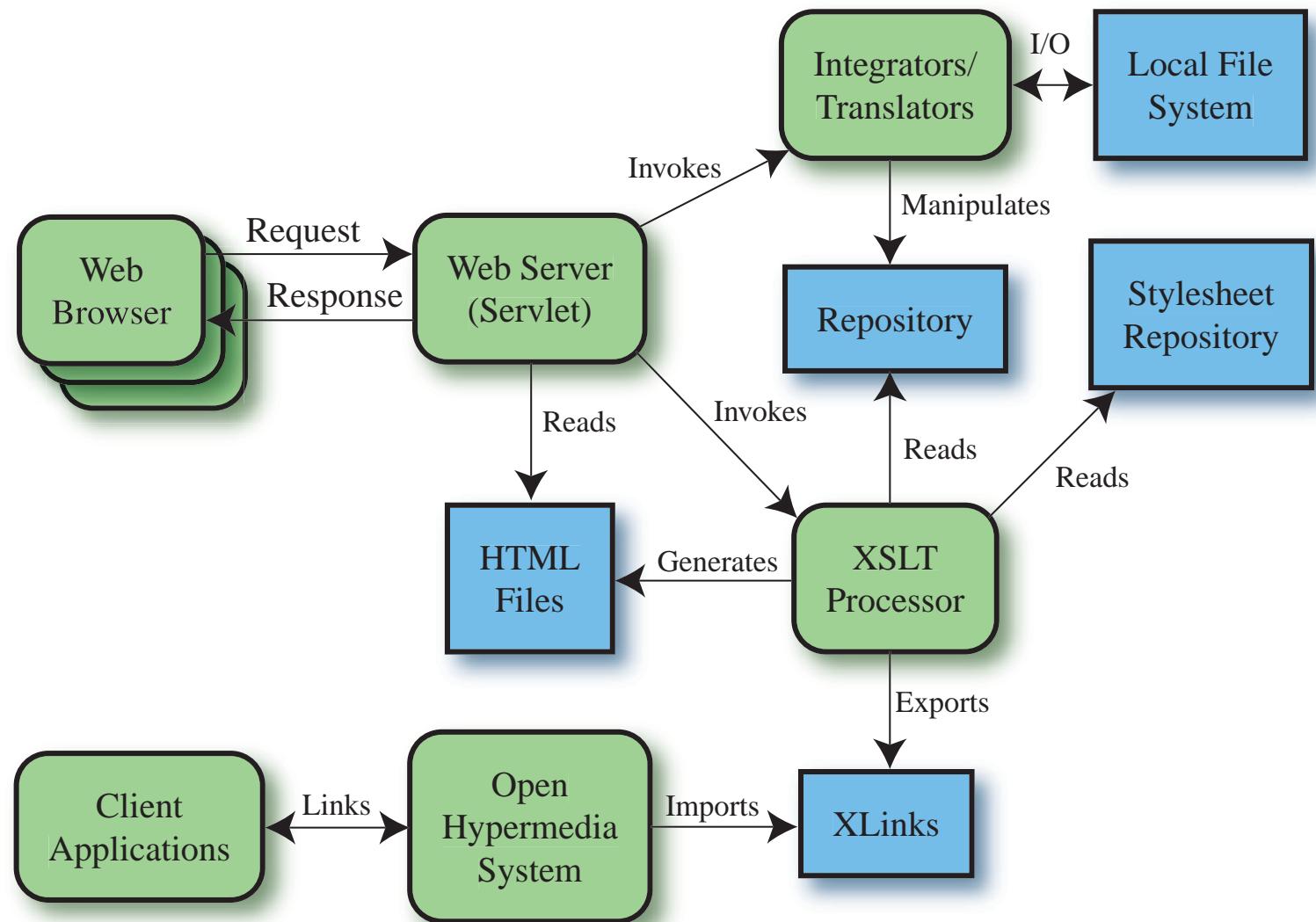
Information Integration

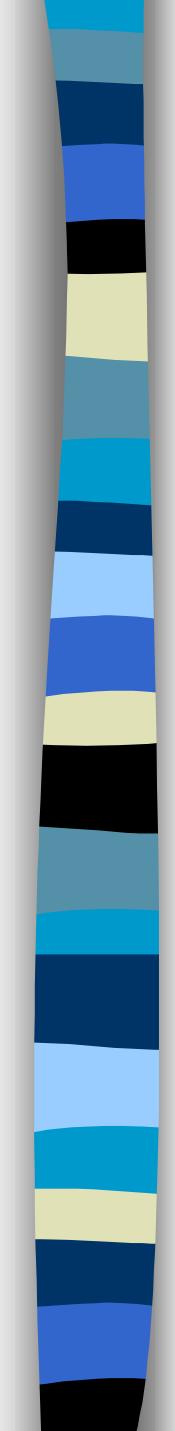
- Many definitions available
 - databases: “schema merge”
 - intelligence community: “information analysis”
- Our definition
 - The task of discovering, creating, managing, and evolving software artifact relationships
- Our approach
 - Building new infrastructure that leverages open hypermedia techniques and technology

InfiniT e Information Integration Environment



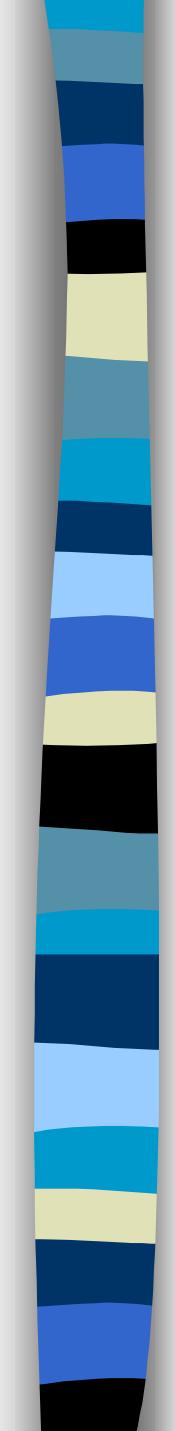
Initial Prototype





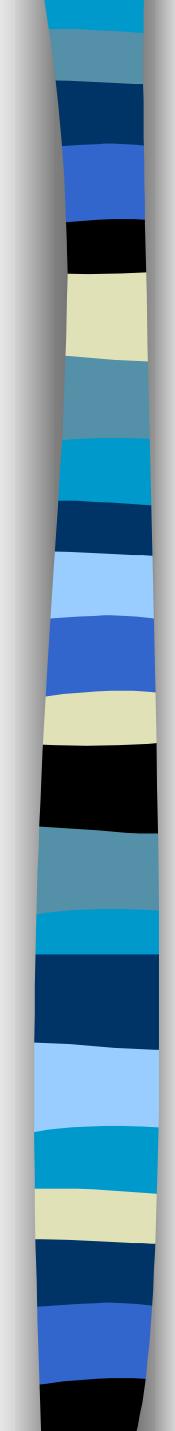
Open Hypermedia Issues

- InfiniTe's design is heavily influenced by the experience of open hypermedia systems
 - Support for heterogeneity
 - Multiple artifact types via multiple translators
 - Note: this design principle is often vigorously resisted!
 - Most recent experience: XML workshop at ICSE
 - Use of open hypermedia concepts
 - InfiniTe contexts are based on composites
 - InfiniTe supports creation of *n*-ary relationships
 - Rest of the world stuck on binary relationships!



Open Hypermedia Issues, cont.

- Results displayed in Native Environments
 - One key benefit of open hypermedia for software engineering is the ability to use “real-world” tools
 - Therefore, InfiniTe is designed to be integrated with an open hypermedia system
 - such that relationships created in InfiniTe can be viewed and traversed over the original software artifacts with an engineer’s favorite set of (open hypermedia aware) tools

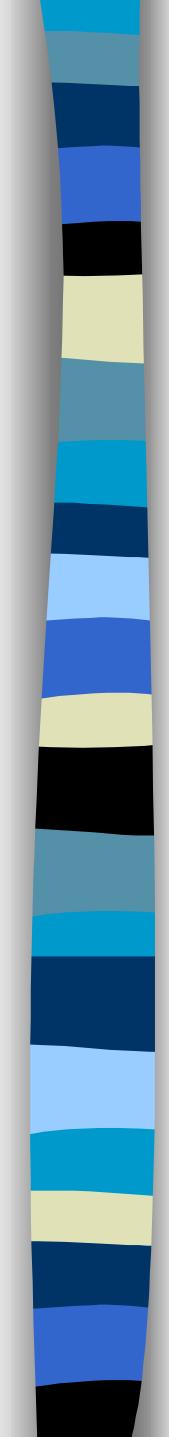


Preconditions for OHS Integration

- InfiniTe keeps track of the original location for all translated documents
- A mapping must exist between XPointers in InfiniTe and LocSpecs in an open hypermedia system that refer to the original document
 - The paper describes an initial mapping used for simple text documents; this mapping is slightly out-of-date with the current implementation and will be updated before final publication

Demo

- Translate two text documents into InfiniT e
- Perform a keyword search
- Export InfiniT e relationships as XLinks
 - Guided Tour Links
 - Index Links
- Import InfiniT e relationships into Chimera
- View relationships in a Chimera text editor



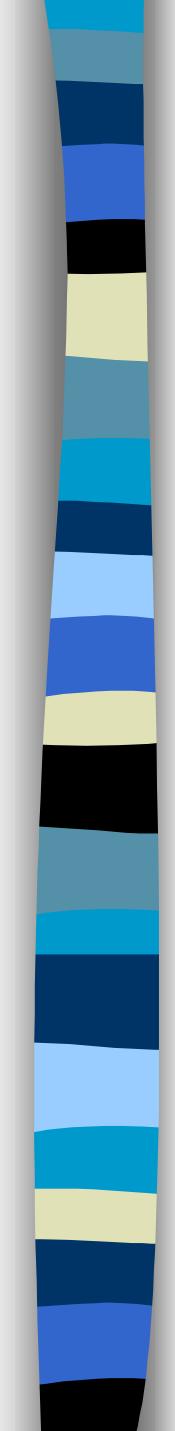
Related Work

■ GeoWorlds

- Intelligence community tool focusing on Web-based data sources and info. analysis techniques
 - relationships are not first class

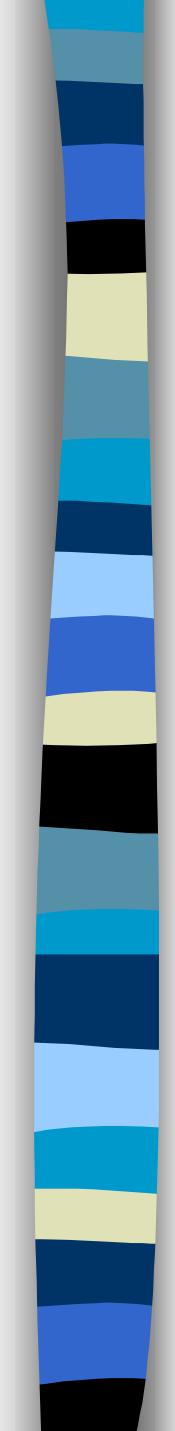
■ xlinkit.com

- Tool for analyzing consistency relationships
 - similar approach but ultimately different goals



Future Work

- More translators
 - Current implementation supports only text documents
- More integrators
 - Need to be able to find more types of relationships
 - Add support for managing sets of relationships over time
- Open Source as Input
 - Use the software artifacts of open source projects as “free data”; generate results and provide them to the open source community for feedback



Conclusions

- Open hypermedia is mature enough to
 - influence the design of systems in external domains
 - provide key linkages with “real-world” tools to serve as a high-fidelity delivery mechanism
 - as opposed to forcing engineers to read and parse XML documents directly!