Reuse in the NSDL

Tom Wrensch, Robby Robson
Eduworks Corporation
Overview

- Questions
- Stories versus Statistics
- Data sources and Data gathering
- Metrics and Tools
- Reuse and Users
- Summary
Types of Reuse

- Adoption (Reuse)
  - Link to a learning object
  - Duplicate on another system

- Adaptation (Repurpose)
  - Modify or extract parts
Questions

• Research Questions
  – How much reuse and repurposing of NSDL is there?
  – What impact is that reuse having?

• Investigation Questions
  – How to detect if an NSDL resource is reused?
  – How to detect if an NSDL resource is repurposed?
  – How to discover the impact of reuse?
Stories vs Statistics

• Lots of statistical information generated
• But...
  – Statistics can be boring, especially in a talk
  – They'll be published in a paper anyway
• Stories are more interesting

• (And maybe just a few statistics)
Metrics and Tools

- Equality via hash values
- Adaptation requires similarity measures
  - Signature extraction (Finkel et al, 2001)
  - Used in plagiarism detection
  - Similar approach for images
- Human focused data via surveys and interviews.
Data Sources and Gathering

- Last year SDSC poster
  - Bing Zhu and Regan Moore
  - Archived the NSDL (yeah, all of it)
  - Nearly a million assets and terabytes of data
Adopting Objects

• Within the NSDL, not much
  – Some small images:
  – Web infrastructure files (CSS, web page templates)

• From the outside into the NSDL, more
  – Mostly linked to rather than copied
  – Curriki: approx 0.02% of links are to NSDL
  – IA: 1.5% of links are to NSDL
Adapting Object

- Interesting results, still partial data
- Within the NSDL
  - Of 3588 word docs 381 pairs are 80%+ similar
  - Most are agreements, notices, and such
- From outside the NSDL
  - Only a few documents, but hints of paragraph-level adaptation.
  - Many images show a high level of similarity
Highly Reusable Objects

• According to existing data
  – Small, simple
  – HTML, images, PDFs
  – Broadly applicable
  – Meaningful in multiple contexts

• But...
Users

• Some surprises
  – 2/3 do not use an LMS
  – More look to learn from than use
  – Pull out paragraphs of text
  – Search weekly or more for learning objects
  – Video clips and simulations are most interesting
  – Appeal to multiple learning styles
  – Worry a lot about copyright and IP issues
Summary

- More data to churn through
- Only a little adoption happening
- More adaptation, especially at find granularity
- No single profile for highly reusable objects
- Looking for different things than in the past
- “Relearning” is important (Assimilate)
- Make intellectual property issues clear
Acknowledgments

- San Diego Supercomputing Center
- NSDL Pathways where critical partners
  - comPADRE
  - MathDL
  - Engineering Pathway
  - CSERD
  - MAT DL
- Work done under NSF grant #033590