Designing Storage Architectures for Preservation Collections

Neighborhood Watch for Repository Quality Assurance

Stephen Abrams
Patricia Cruse
John Kunze

University of California Curation Center
California Digital Library

David Minor

San Diego Supercomputer Center
University of California, San Diego

Mike Smorul

University of Maryland Institute for Advanced Computer Studies

Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple
Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple
Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple

Lather, rinse, repeat...

http://www.flickr.com/photos/yenzah/5017275397
Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple
Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple

Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple

• What is the fundamental obligation of a repository?
  – First, do no harm [primum non necere]
Neighborhood watch

• A new metric, and community ethic, for repository reliability and trustworthiness
  – Objective
  – Repeatable
  – Independently verifiable
  – Simple

• What is the fundamental obligation of a repository?
  – First, save the bits
    ... before worrying about added-value services
Neighborhood watch

• Based on information that is already readily available
  – Content retrieval via stable URLs
  – Content size and message digest
Neighborhood watch

• Based on information that is already readily available
  – Content retrieval via stable URLs
  – Content size and message digest

• An external agent that periodically retrieves content, verifies its bit-level integrity, and notifies the neighborhood

• Possible implementations
  – Merritt Fixity service
    http://www.cdlib.org/uc3/curation/fixity.html
  – UMIACS ACE
    http://adapt.umiacs.umd.edu/ace
How to be a good neighbor

• “Have an open door policy”
  – Repositories should provide access to all content

• “Don’t borrow too much sugar”
  – Verification agents should throttle requests

• “If you see something, say something”
  – Adhere to community standards for reporting and veracity
So …

- Trust, but verify
- What you say about yourself is less interesting than what others say about you
- Be neighborly
So …

• Trust, but verify

• What you say about yourself is less interesting than what others say about you

• Be neighborly
So …

• Trust, but verify

• What you say about yourself is less interesting than what others say about you

• Be neighborly
So ...

• Trust, but verify

• What you say about yourself is less interesting than what others say about you

• Be neighborly


http://www.flickr.com/photos/48304881@N05/5240756741

National Neighborhood Watch Institute