NEW DYNAMICS CREATE NEW IDEAS

The First Year of Action by the National Digital Stewardship Alliance

The National Digital Stewardship Alliance's 100 + members are a diverse group of universities, commercial businesses, professional associations, non-profit and government agencies at the federal, state, and local level.

For information about the NDSA visit
http://www.digitalpreservation.gov/ndsa/

THE WORK OF THE GROUPS

Content: Opt-In to Web Archiving: Developing the Preservation Plug-in

Innovation: "Neighborhood Watch" for Repository Quality Assurance

Infrastructure: How We Do Storage? A Saga of Bits, Disk, Tape and Systems

Standards: Standards-ipedia: Adding Digital Preservation Information to Wikipedia

Outreach: Dataloss Stories: Establishing Urgency through Storytelling

OPT-IN TO WEB ARCHIVING

Developing the Preservation Plug-in

CONTENT WORKING GROUP

Focus:

- Identifying content already preserved
- Investigating guidelines for the selection of significant content,
- Discovery of at-risk digital content or collections, and
- Matching orphan content with NDSA partners who will acquire the content, preserve it, and provide access to it.

Active Projects:

- Guidance on appraisal for local government information.
- A blog preservation plugin to easily enable content creators to opt-in to preservation.
- A survey of organizations in the United States who are actively involved in or planning to archive content from the web.

PRESERVATION PLUG-IN

Goals: Active web preservation initiated by content publishers

- Critical need for simple mechanisms to Opt-in to Web collection building
- Built matrix identifying key publishing platforms to approach (ex. WordPress, Typepad, Squarespace, Live Journal, Blogger, Xanga, OpenSalon, Drupal, etc.)
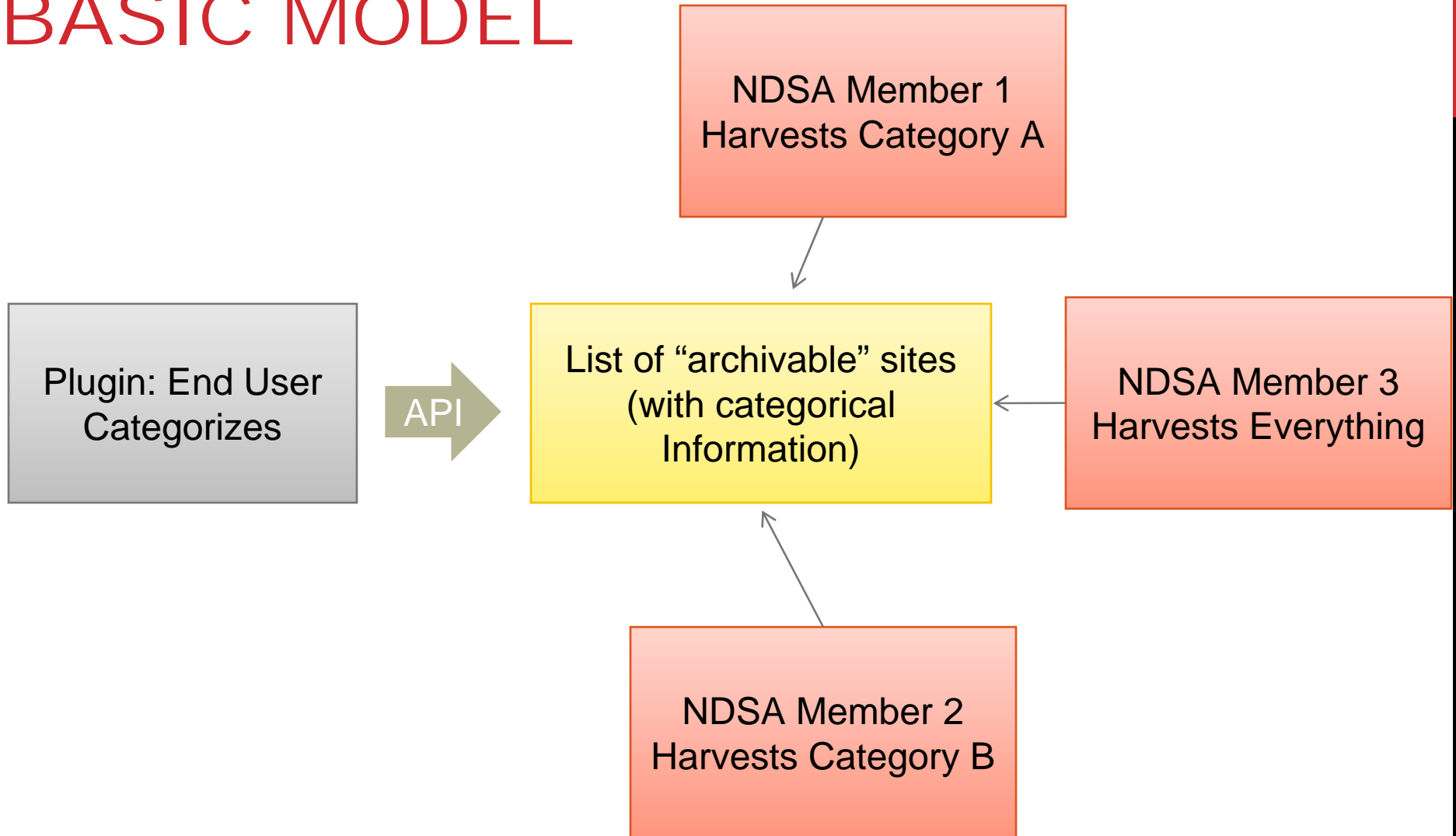
Launching Pilot w/WordPress

- Integrating w/ downloadable OS software
- Enabling site publisher to Opt-in and later view which NDSA institutions have selected their site for inclusion in collection/s
- NDSA partners can view publisher submissions and mark those matching their selection criteria, also indicating when the site will be/was captured
- IA commits to collect all submissions at least once per year so site publishers are guaranteed inclusion in at least one archive

Timeline

- January 2011: Build plug-in and supporting services
- February 2012: Deploy to site publishers as optional plug-in

BASIC MODEL

NDSA Member 1
Harvests Category A

Plugin: End User
Categorizes

API

List of "archivable" sites
(with categorical
Information)

NDSA Member 3
Harvests Everything

NDSA Member 2
Harvests Category B

"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

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

NEIGHBORHOOD WATCH

A new metric, and community ethic, for repository reliability and trustworthiness
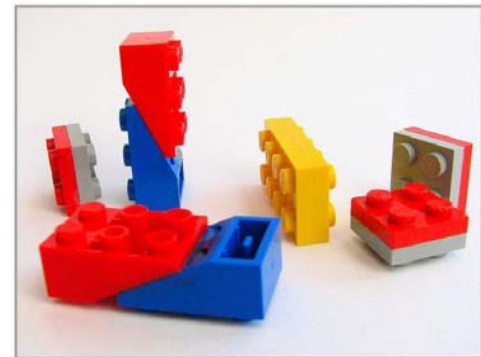
- Objective

- Repeatable

- Independently verifiable

- Simple

NEIGHBORHOOD WATCH

A new metric, and community ethic, for repository reliability and trustworthiness

•Objective

•Repeatable

•Independently verifiable

•Simple

WindellOskay, http://www.flickr.com/photos/oskay/265899865

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..

1. Trust, but verify

2. What you say about yourself is less interesting than what others say about you

3. Be neighborly

SO..

1. Trust, but verify

2. What you say about yourself is less interesting than what others say about you

3. Be neighborly

AFP/Getty Images, http://www.britannica.com/bps/media-view/121436/1/0/0

SO..

1. Trust, but verify

2. What you say about yourself is less interesting than what others say about you

3. Be neighborly

SO..

1. Trust, but verify

2. What you say about yourself is less interesting than what others say about you

3. Be neighborly

HOW WE DO STORAGE?

A Saga of Bits, Disk, Tape and Systems

INFRASTRUCTURE WORKING GROUP

Purpose: to build a community from sharing information and best practices about the development and maintenance of tools and systems for the curation, preservation, storage, hosting, migration, and similar activities for the long term preservation of digital content.

Goal: to survey and document best practices in the digital preservation infrastructure.

THE EVOLUTION OF A PROJECT…

Original interest in cloud storage for preservation

- Interviewed cloud providers
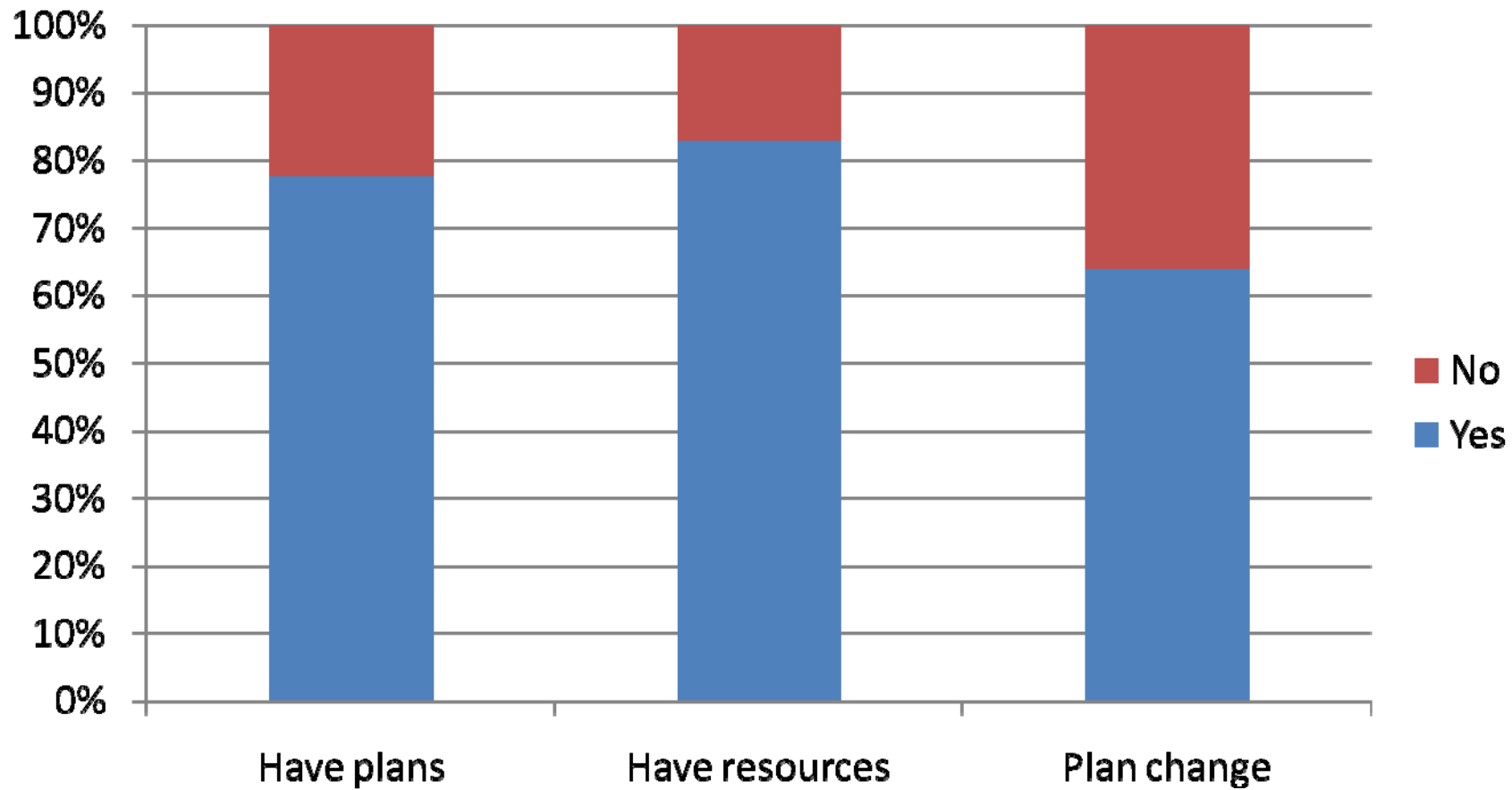- Built matrix comparing 7 services

13 open-ended questions for Working Group members

- How they do/do not use cloud storage
- Other storage-related practices and concerns
- Summarized 10 responses in storage report
- Presented at NDSA Making It Work meeting
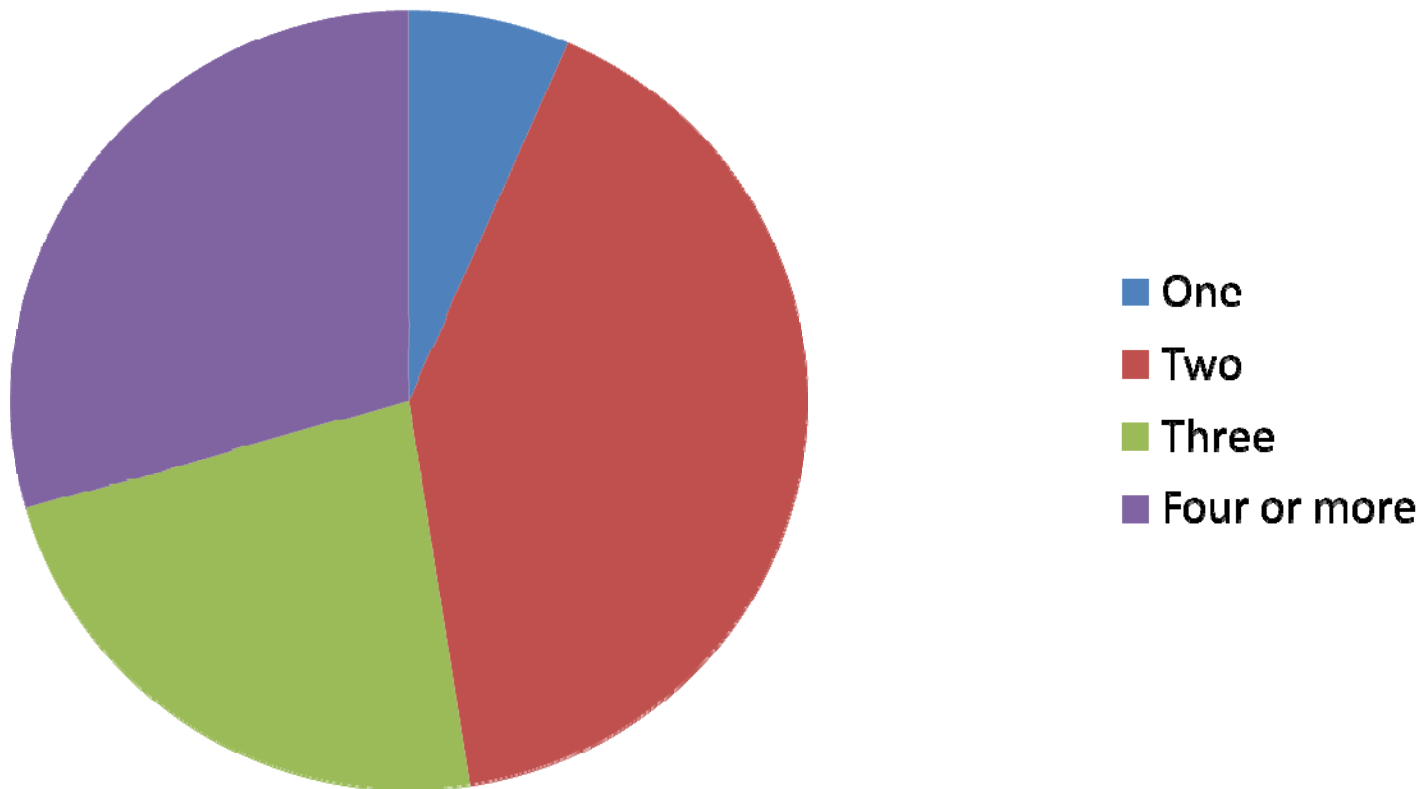
Larger, more quantitative survey for all NDSA members

- 21 questions
- 58 responses (78% response rate)
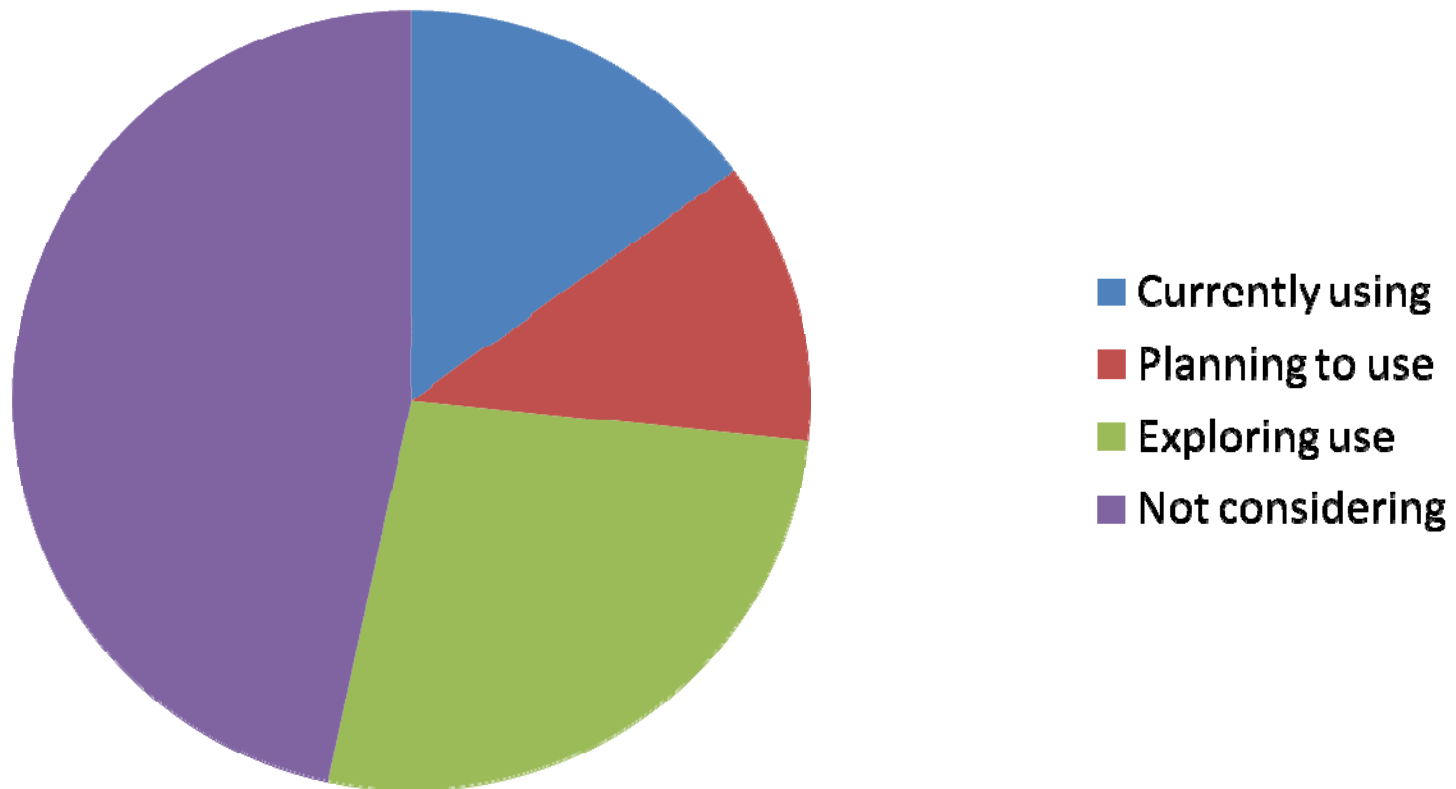- Examples of results:

THREE YEAR PLANS

NUMBER OF COPIES KEPT

member institutions

Legend:
- One
- Two
- Three
- Four or more

USE OF CLOUD STORAGE

Member organizations

- Currently using
- Planning to use
- Exploring use
- Not considering

FUTURE PLANS

Follow up and deeper analysis of survey results

- Are there trends by institution size or type?
- Can some apparent inconsistencies be resolved by email?
- Why did non-respondents not reply?

Integrate results into original storage report

Consider a shorter survey of the entire community (beyond NDSA)

Next up: Guidance for Open Source Software providers and adopters

STANDARDS-IPEDIA

Adding Digital Preservation Information to Wikipedia

Wikipedia has become the go-to resource for technical information on things like file formats and standards. Instead of creating another resource on digital preservation standards, the NDSA standards working group is focusing on enhancing coverage and sophistication of information and resources on born digital preservation concerns and criteria in Wikipedia.

DATALOSS STORIES

Establishing Urgency through Storytelling

National Digital Stewardship Alliance: Request for stories of data loss and/or preservation

Have you, or has your institution, lost important data through 1) hardware failure, 2) hardware or format obsolescence, 3) migration to a new format, or other means? Were you once in danger of losing important data, but you managed to recover or preserve it? The National Digital Stewardship Alliance would love to hear about it --

For information about the NDSA visit
http://www.digitalpreservation.gov/ndsa/

To join the NDSA visit
http://www.digitalpreservation.gov/ndsaform/