



STANFORD UNIVERSITY LIBRARIES

Outreach to Campus Webmasters for a Better Web, and Better Web Archiving

Nicholas Taylor

<u>Web Archiving</u> Service Manager

<u>Stanford University Libraries</u>

Archives 2015 Session 306 – Seeding Engagement: Web Archiving Outreach Opportunities and Strategies August 21, 2015



webmasters don't care about web archiving





webmasters more concerned w/ Web-mastery



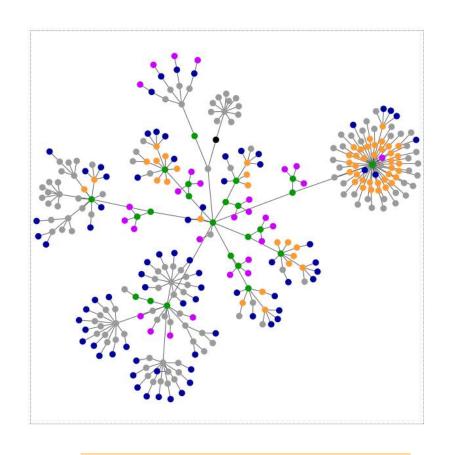
look for points of alignment





decentralized Web production

- Stanford Web Services
- Stanford Communications
- tens of distributed webmasters for administrative units
- many more individual content creators
- no central coordination
- centralized services available (Drupal, AFS) but diversity of platforms used



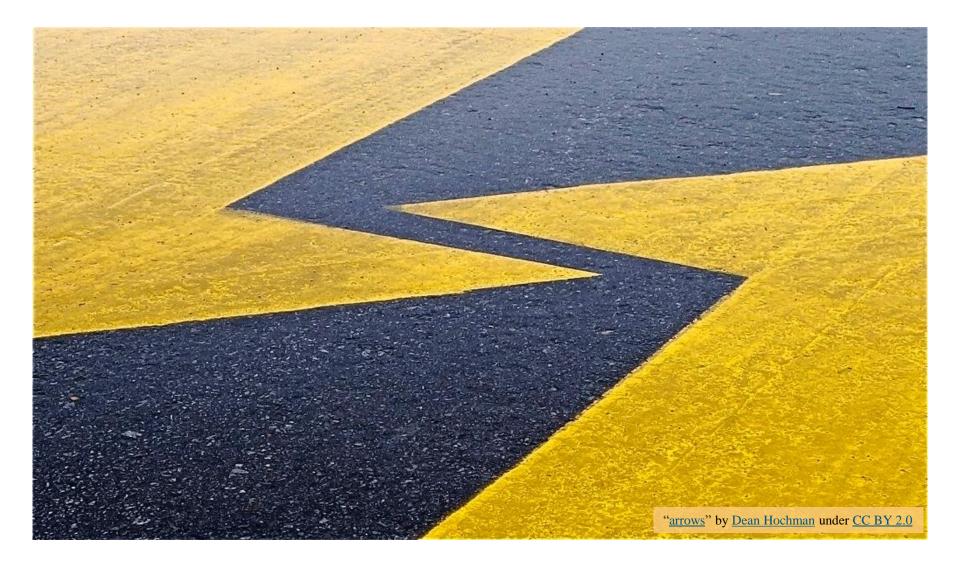


become a free SEO consultant



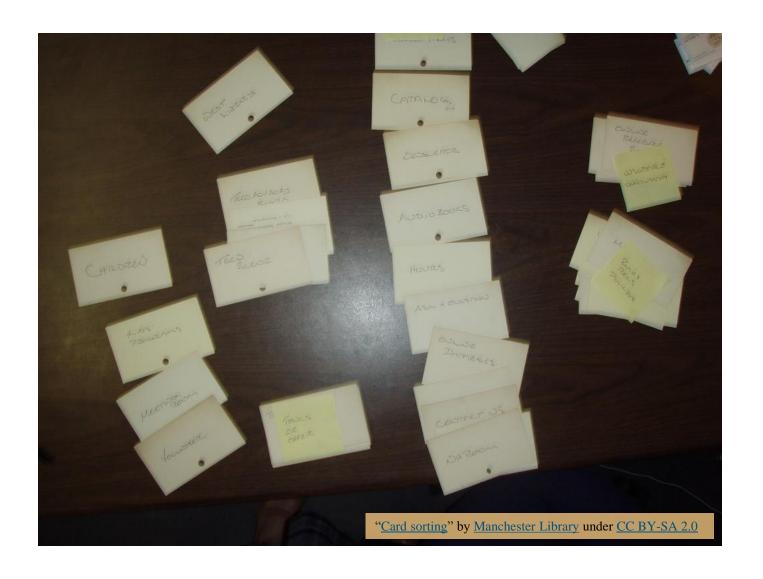


discoverable links





planned site structure





responsive design



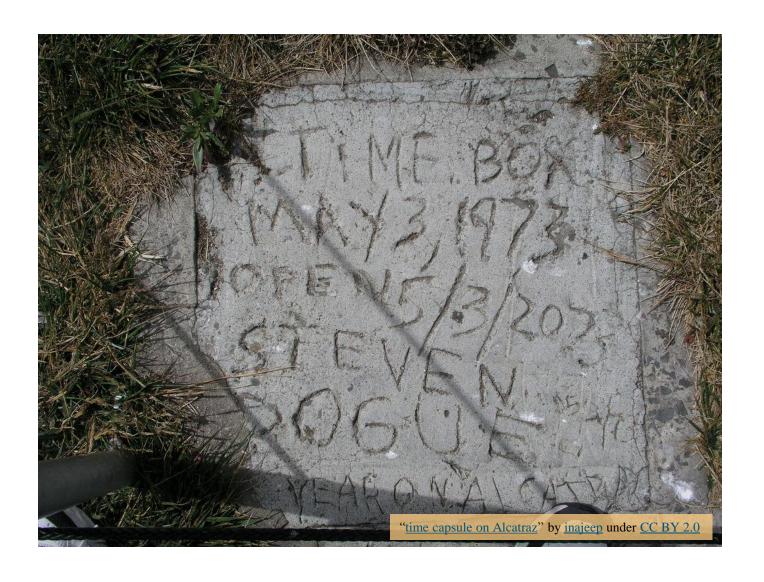


reusable assets



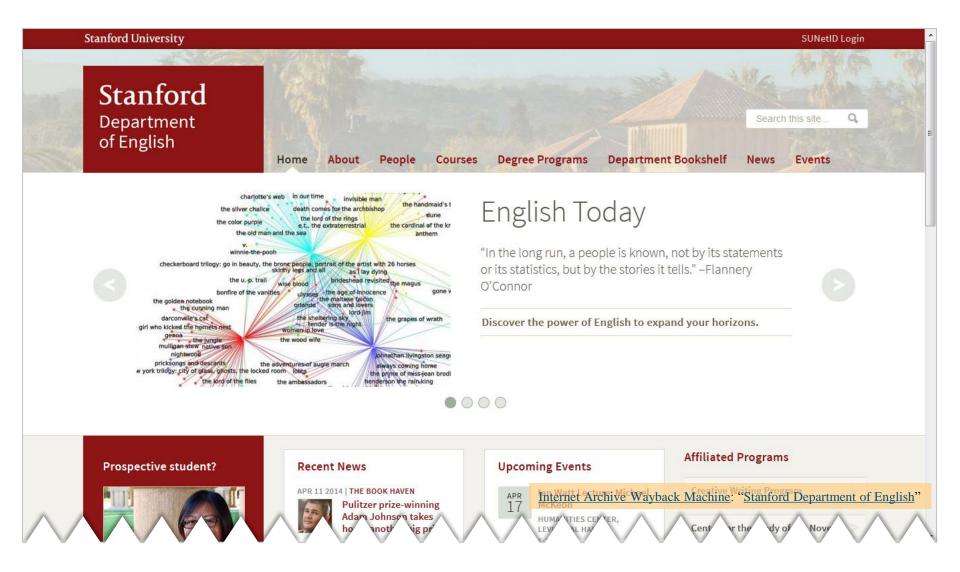


caching headers



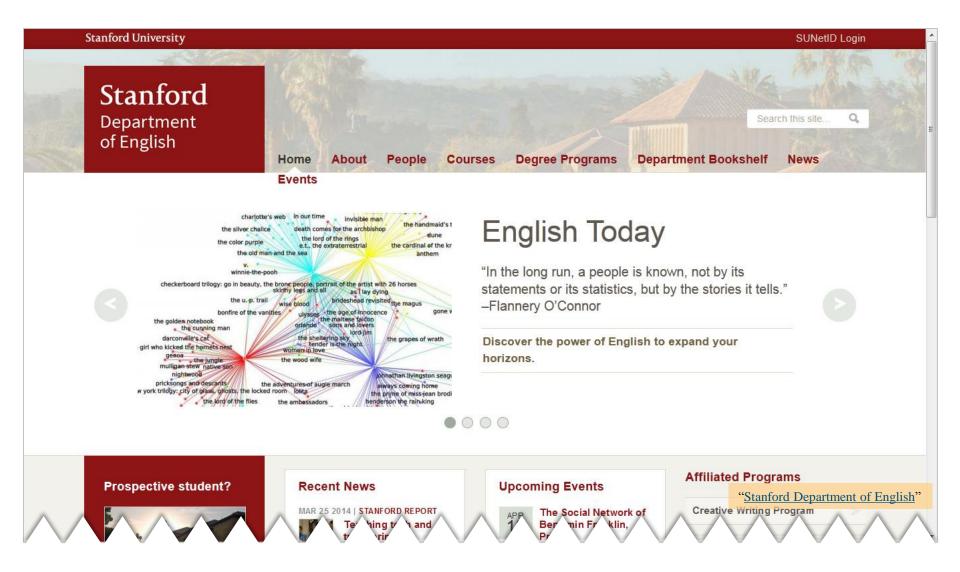


reduce reliance on external assets





reduce reliance on external assets



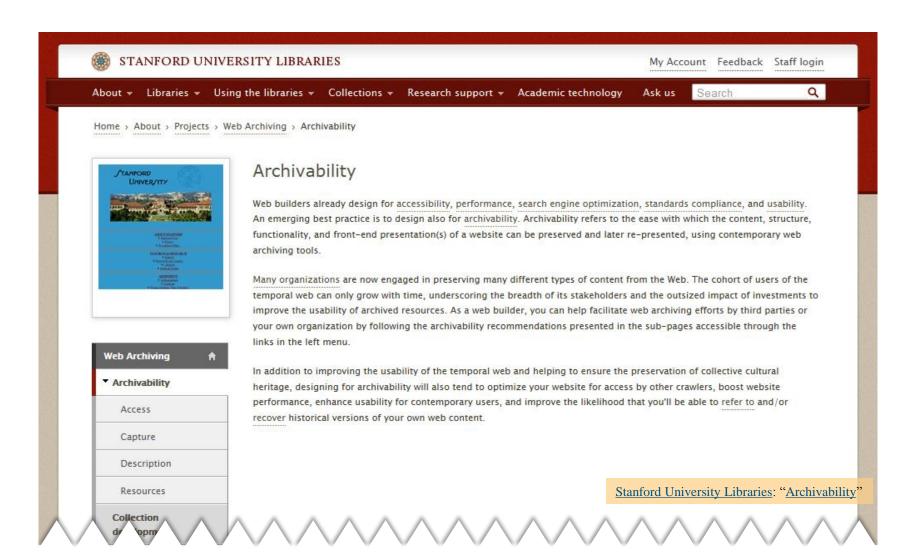


partnership opportunities



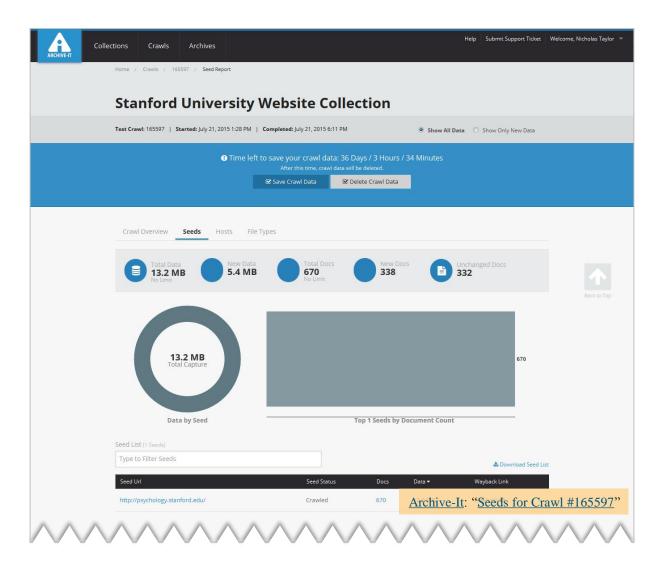


archivability guidance



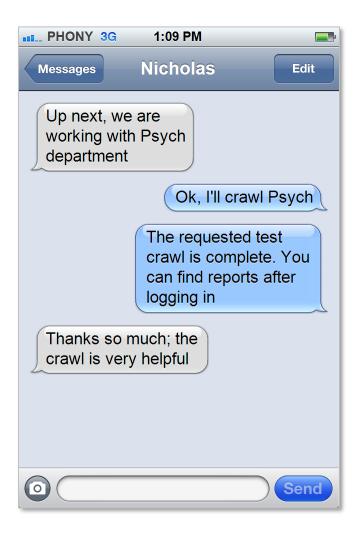


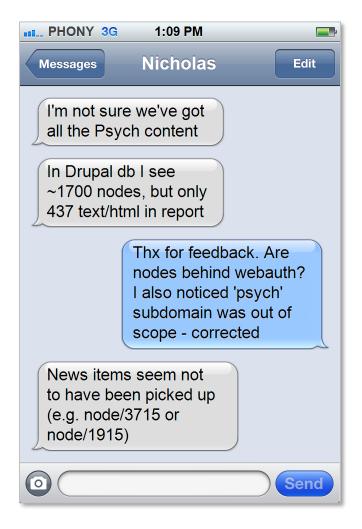
test crawls for content assessment





more complete crawls







locate deprecated but archived content



STANFORD UNIVERSITY LIBRARIES

Stanford Web Archive Portal Home | Help

Showing http://slacvm.slac.stanford.edu/FIND/slac.html captured on Oct 4, 1994

♦ Previous capture | Show overlay | Next capture >



WorldWideWeb SLAC Home Page

SLAC 7 Sep 1994

Use the WorldWideWeb (WWW) service to gain access to a wide range of information at SLAC and elsewhere around the globe. Emphasized text like this is a hypertext link.

You may view WWW information through GUI or line-mode browsers. At least most SLAC pages have been tested on the MidasWWW X Window System browser.

Note that over time links may move around on a page, migrate to other pages, or be removed entirely, due to the dynamic nature of the Web. That there is no current Hot Topic illustrates this characteristic today.

SLAC Information

People and organizations:

people at SLAC, particle physics people and institutions; ESnet white pages.

Library:

SPIRES-HEP, Current PPF-list, Books, SLACspeak glossary, Library News, more databases.

Physics Preprint Bulletin Boards (full-text postscript):

today, yesterday, last seven days, week before that, let me search.

Seminars:

today, tomorrow, this week, next week, this year, let me search.

Conferences

this month, next month, next summer, next year, all future, let me search.

News:

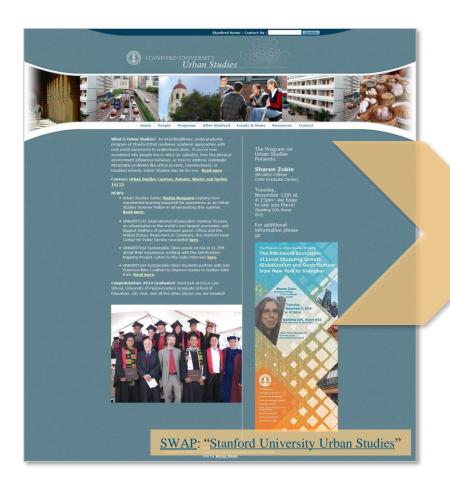
APS What's New, SLAC Netnews.

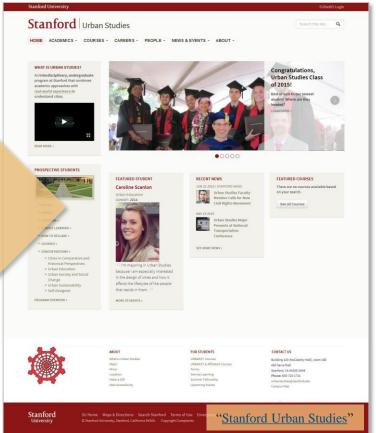
SWAP: "WorldWideWeb SLAC Home Page"

St AC Physic



take timely snapshots





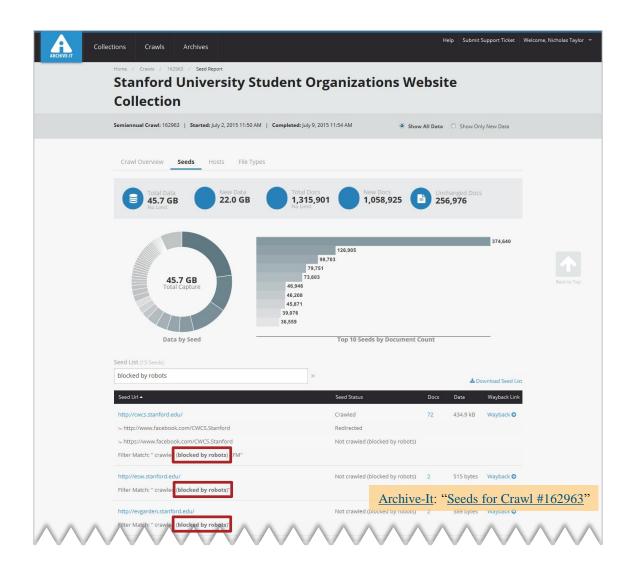


Archive-It crawler isn't malicious





contacts to unblock robots



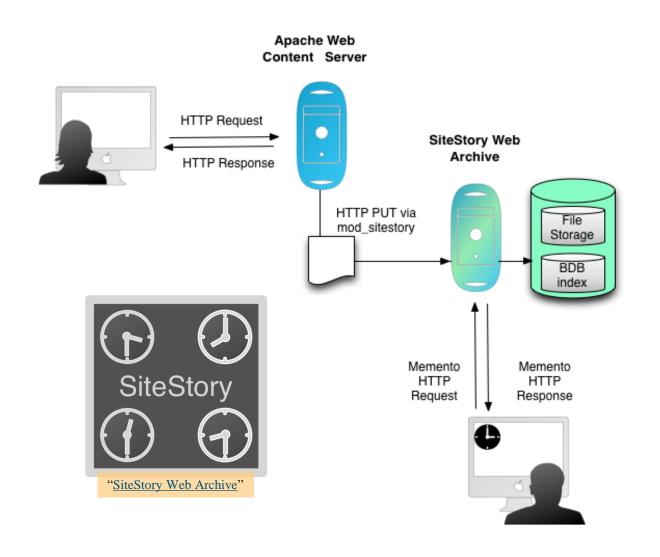


custom 404 page

```
(function() {
    var s = document.createElement('script');
                                                                                              Would you like to see an archived version of this page in the
    s.type = 'text/javascript';
                                                                                              Internet Archive's Wayback Machine?
    s.asvnc = true;
    var page url = window.location.href;
    page_url = page_url.replace('//www-ilya.', '//'); //strip
    page url = page price
    var wb404 url = https://archive.org/wayback/available.php
                                                              callback=wb404_callback&url='+page_url;
    s.src = wb404 ur
    var h = document.getElementsByTagName('head').item(0) || document.documentElement;
    h.appendChild(s);
    var l = document.createElement('link');
    1.rel = 'stylesheet';
    1.type = 'text/css';
                                                                            https://archive.org/wayback/available.php
    1.href = 'https://archive.org/web/wb404.css';
    h.appendChild(1);
})();
wb404 callback = function(obj) {
    var archived text = "Would you like to <a href='$url' onClick='wb404 record click(this); return false;'>see an archived version of this page</a> in the Internet
Archive's Wayback Machine?";
    var maybe text = "Would you like to <a href='$url' onClick='wb404 record click(this); return false;'>check the Internet Archive's Wayback Machine</a> for an
archived version of this page?";
    var wb image = 'https://archive.org/images/wayback404.png';
    if (!obj.archived snapshots || !obj.archived snapshots.closest || !obj.archived snapshots.closest.available) {
       return false:
    var url = obj.archived snapshots.closest.url;
    var html = "<br/>div class='wb404 imagediv'><a href='"+url+"' onClick='wb404 record click(this); return false;'><img class='wb404 image' src='" + wb image +
"'/></a></div>";
    if (true) {
       html += "<div class='wb404 text'>" + archived text.replace('$url', url) + "</div><br clear='both'/>";
        //Not supporting this for now
        html += "<div class='wb404 text'>" + maybe text.replace('$url', url) + "</div><br/>br clear='both'/>";
    var wb404 div = document.getElementById('wb404');
    wb404 div.innerHTML = html;
                                                              Internet Archive: "Free "404: File Not Found" Handler for Webmasters to Improve User Experience"
wb404 record click = function(link) {
    var img = new Image(1,1);
```



transactional web archiving





webmasters don't care about web archiving





webmasters may care about web archiving

