# TECHNIQUES FOR MANAGING EVOLVING WEB CONTENT

SEAN BROWN
IS&T - WEB COMMUNICATION SERVICES

< SMBROWN@MIT.EDU

HTTP://WEB.MIT.EDU/IST/WEB/

### **OVERVIEW**

- Introduction
- Outline of Presentation
- Audience Survey

### TECHNIQUES FOR MANAGING EVOLVING WEB CONTENT

- Making site-wide updates easier
- What to do when you need to move pages around
- How to manage major changes, like launching a new version of your site
- Give your audience a better user experience when things go wrong
- Improve search engine effectiveness

### PRIMARY AUDIENCE

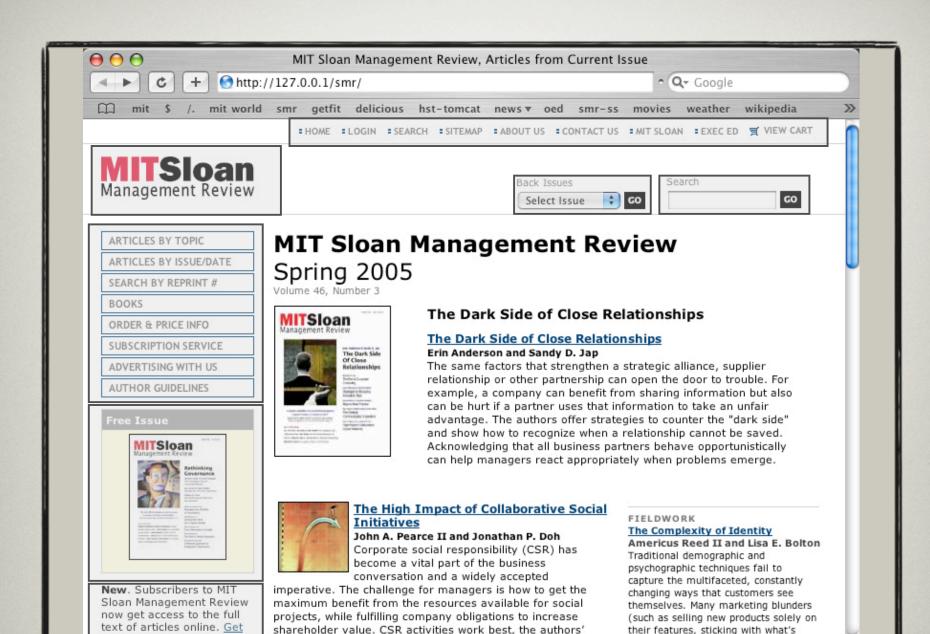
- Web publisher at MIT
- Medium large website
- Primarily static website
- Use the source Luke
- web.mit.edu

### TOPICS COVERED

- Maximizing content re-use with some advanced server-side include (SSI) tricks
- What to do when your site starts changing;
   doing content re-direction the Right way
- Custom error pages; they won't thank you, but they should
- Search engine impact, plus a new twist on an old idea

### SERVER-SIDE INCLUDES

- Most often used for content that repeats verbatim
- Did you know that it can be used for variable content?
  - Rotating news, callouts, special events
  - Site navigation (including highlight states)
  - XBitHack When you don't want to .shtml



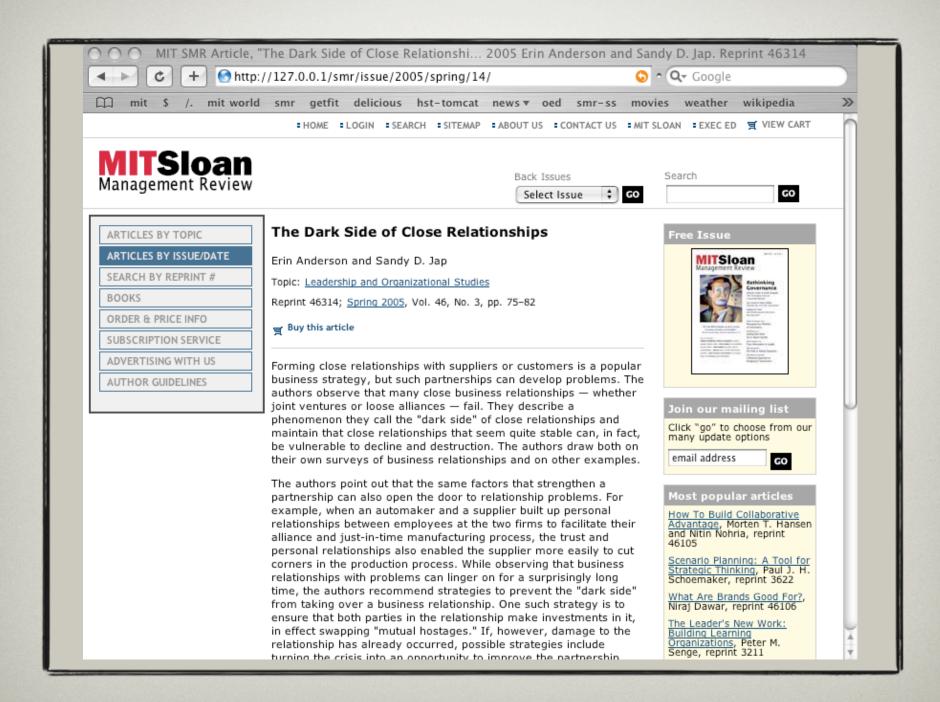
research shows, when they are managed strategically and

are conducted in collaboration with an array of commercial

worked before and underestimating

low-involvement products) can be

more information.



# How can one SSI be used for nav element with multiple states?



### In each page, set template type:

#### Using plain HTML, the code looks like this:

```
<div class="nav">
    <a href="/smr/topic/">Articles by Topic</a>
</div>
Kdiv class="highlight_nay">
    <a href="/smr/issue/">Articles by Issue/Date</a></a>
</div>
<div class="nav">
    <a href="/smr/search/">Search by Reprint #</a>
</div>
<div class="nav">
    <a href="/smr/book/">Books</a>
</div>
<div class="nav">
    <a href="/smr/ordering/">Order &amp; Price Info</a>
</div>
Kdiv class="nav">
    <a href="/smr/subscribe/">Subscription Service</a>
</div>
<div class="nav">
    <a href="/smr/advertise/">Advertising with Us</a></a>
</div>
<div class="nav">
    <a href="/smr/authoring/">Author Guidelines</a></a>
</ta>
```

#### SSI's can do conditional evaluations, like these:

```
<!--#if expr="${template} == \"topic\"" -->
<div class="highlight_nav">
<!--#else -->
Kdiv class="nav">
<!--#endif -->
    <a href="/smr/topic/">Articles by Topic</a></a>
\langle !--#if expr="${template} == \"issue\"" -->
<div class="highlight_nav">
<!--#else -->
<div class="nav">
<!--#endif -->
    <a href="/smr/issue/">Articles by Issue/Date</a>
\langle !--#if expr="${template} == \"search\"" -->
<div class="highlight_nav">
<!--#else -->
Kdiv class="nav">
<!--#endif -->
<!-- ... continues ... -->
```

### CONTENT REDIRECTION

- Why we do it in the first place
  - Link rot, it's evil
  - Link rot effects search engine placement
  - Breaks the user experience
- The oft-used "Meta-Refesh" Hack

### META-REFRESH IS BAD

- (Most of the time)
- About HTTP Response Headers
- Temporarily Moved (302) vs. Moved Permanently (301)
- Cumulative effect. Have you ever followed three "this page has moved…" or more pages only to wind up on a 404 / Document Not Found page? I have.
- Search engine impact

### META-REFRESH - CONT.

- When it's OK
  - Content truly has moved temporarily (will be back < 1 month)
  - Landing pages (tracking campaigns, etc.)
  - Slide-shows, transition pages
- But you can use mod\_alias to do the same thing, without creating pages, or delays

# WHAT ELSE CAN I DO THOUGH?

- mod\_alias Configurable in .htaccess.mit files (Site-wide, per directory)
- Apache directive that provides two components: content aliasing, content redirection
- Can use regular expressions to match patterns in URLs; Let's you redirect lots of content with one or a few statements

## CONTENT REDIRECTION-EXAMPLE

 You've moved all of the content from /wcs/conference/
 -to-/wcs/conference/05/

• in an .htaccess.mit file in the root level place this:

Redirect 302 / wcs/conference(.\*)\$
<a href="http://web.mit.edu/wcs/conference/05/\$1">http://web.mit.edu/wcs/conference/05/\$1</a>

### MORE SCENARIOS

- Converting all your .GIFs to .JPG/PNG
- Moving all your content from one Athena locker to another (or to a dedicated server)
- Site redesigns usually mean new Information Architectures, mod\_alias' Redirect and RedirectMatch can help

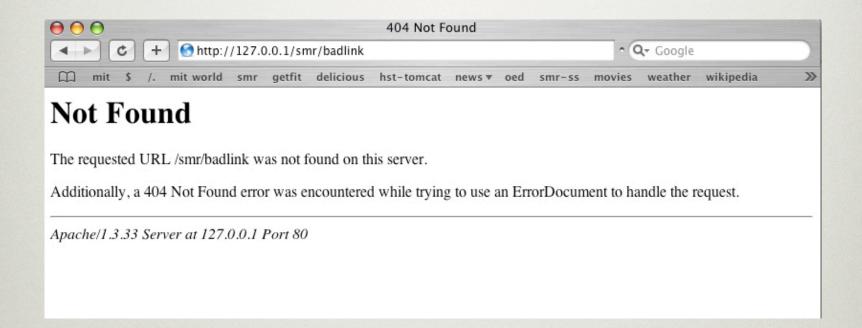
### **ALIASING CONTENT**

- Like a Shortcut / Symlink / Alias but for URLs
- Can be used to combine Athena lockers into one virtual site
- Also useful sometimes if you need to reuse some assets but they need to look like they are in different places.

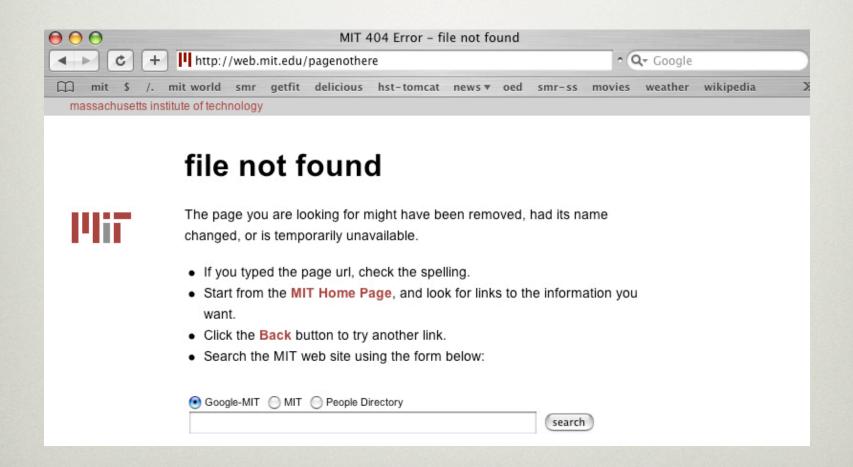
# CUSTOM ERROR DOCUMENTS

- Most popular: custom 404 pages
- Provide branded, consistent interface when errors happen (document not found, SSL required, server error)
- Offer search, links to popular resources

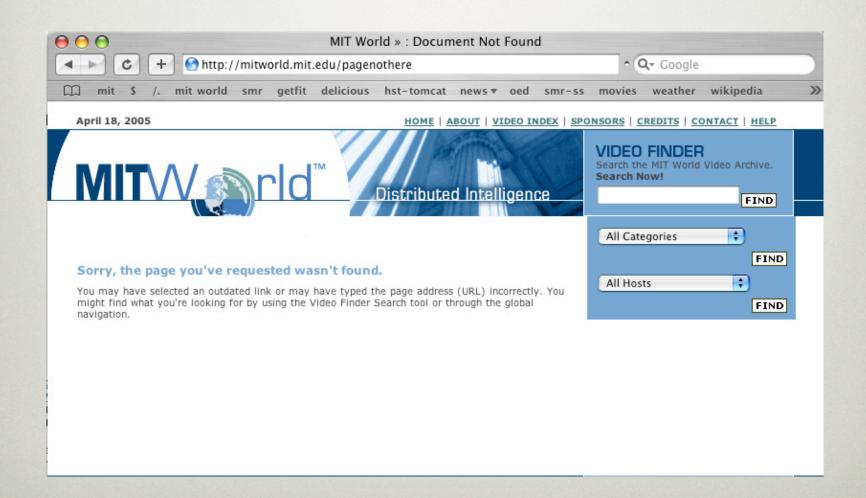
### BASIC 404 PAGE



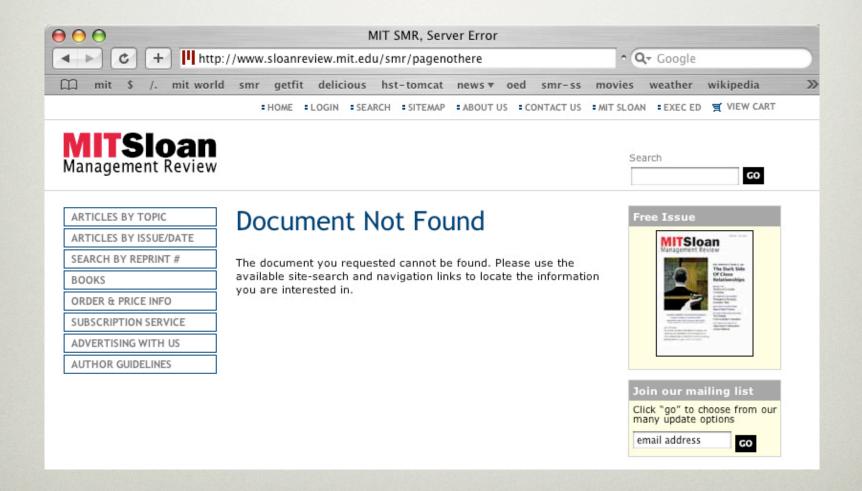
### WEB.MIT.EDU 404 PAGE



### CUSTOMIZED 404 PAGE



### CUSTOMIZED 404 PAGE



### SEARCH ENGINES

- Returning a relevant HTTP response code is important
- Meta-Refresh pages confuse search results
- Permanent redirects makes search engine's update their results index
- Link rot interferes with results ranking

### QUESTIONS & ANSWERS

- Thank you
- Sean Brown < smbrown@mit.edu >

### REFERENCES / LINKS

- IS&T WCS Web Reference
  - http://web.mit.edu/ist/web/reference/
- SSI
  - <a href="http://web.mit.edu/ist/web/reference/web-resources/ssi.shtml">http://web.mit.edu/ist/web/reference/web-resources/ssi.shtml</a>
  - http://httpd.apache.org/docs/howto/ssi.html
- Mod\_alias
  - <a href="http://httpd.apache.org/docs/mod/mod\_alias.html">http://httpd.apache.org/docs/mod/mod\_alias.html</a>
- Custom Error Documents
  - <a href="http://httpd.apache.org/docs/mod/core.html">http://httpd.apache.org/docs/mod/core.html</a>
- MIT's Google Search Engine
  - search@mit.edu