a cost-benefit approach for describing and processing digital objects

http://texashistory.unt.edu/

Dreanna Belden
dreanna.belden@unt.edu
August 30, 2008
in a perfect world . . .

you control the vertical

and the horizontal.

do not change the channel.
in the real world...
issues with “preciousizing”

• target vs. neiman’s approach
• perfection ≠ usefulness
• perceived value
• access = gold standard
philosophical questions . . .

• how valid are our determinations of low and high value of collections? what constitutes significance, and to whom?

• who is the audience? (are you sure?)

• uncorrected or corrected OCR?

• full or select transcriptions? XML markup?

• level of description?
a tale of two collections . . .
a tale of two collections . . .
a tale of two collections . . .

<table>
<thead>
<tr>
<th></th>
<th>Lorenzo de Zavala v.</th>
<th>Deaf Smith County</th>
</tr>
</thead>
<tbody>
<tr>
<td># of digital images</td>
<td>1093</td>
<td>3,662</td>
</tr>
<tr>
<td>page views</td>
<td>257,763</td>
<td>523,583</td>
</tr>
<tr>
<td>historical significance</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>time spent/record</td>
<td>a LOT</td>
<td>minimal</td>
</tr>
<tr>
<td>online since</td>
<td>2005</td>
<td>2006</td>
</tr>
<tr>
<td>cost of project</td>
<td>$21,578</td>
<td>$18,581</td>
</tr>
<tr>
<td>cost/page</td>
<td>$19.75</td>
<td>$5.07</td>
</tr>
<tr>
<td>cost/page view</td>
<td>.0837</td>
<td>.0355</td>
</tr>
</tbody>
</table>
a tale of two more collections . . .
a tale of two more collections . . .

<table>
<thead>
<tr>
<th></th>
<th>Annexation of Texas</th>
<th>Soil Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td># of digital images</td>
<td>5353</td>
<td>4823</td>
</tr>
<tr>
<td>page views</td>
<td>667,150</td>
<td>590,765</td>
</tr>
<tr>
<td>significance</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>time spent/record</td>
<td>significant</td>
<td>minimal</td>
</tr>
<tr>
<td>online since</td>
<td>2004</td>
<td>2006</td>
</tr>
<tr>
<td>total cost</td>
<td>$47,187</td>
<td>$7,511</td>
</tr>
<tr>
<td>cost/digital image</td>
<td>$8.81</td>
<td>$1.56</td>
</tr>
<tr>
<td>cost/page view</td>
<td>.0707</td>
<td>.0127</td>
</tr>
</tbody>
</table>
cost benefit analysis graph

Figured as the value of page cost per view over time. The benefit being measured is usage. (cost per digital image in cents/page views per year).
Figured as the value of page cost per view over time. The benefit being measured is usage. (cost per digital image in cents/page views per year).
who is the audience?
what role can automation play?

Texas Instruments Semiconductor Plant, Dallas, Richardson Public Library
what role can automation play?

• automate population of data fields

• ingest data from existing descriptive systems, such as OPAC or other databases

• aim for adequate, minimal, sufficient
systems to manage data entry

meeting adequate, minimal, sufficient

Created by automated means

Created /Enhanced by resource owners

Created/Enhanced by metadata team

UNTLL Metadata Records

Bring on the shovels . . .

automate population of data fields to meet adequate, minimal, sufficient
automate population of data fields to meet adequate, minimal, sufficient
automate population of data fields to meet adequate, minimal, sufficient
automate population of data fields to meet adequate, minimal, sufficient
automate population of data fields to meet adequate, minimal, sufficient
ingest data from existing descriptive systems

• FileMaker Pro
• Access database
• MARC records
• Crosswalks to Dublin Core, MARC, MODS, PREMIS

http://www.library.unt.edu/digitalprojects/metadata/mapping
ingest data from existing descriptive systems
beware the enemy . . .

- most cases ≠ ideal standards
- perceived uniqueness
- sometimes less is more
thank you!

Dreanna Belden
University of North Texas
dreanna.belden@unt.edu