is an allegory used by the Greek philosopher Plato in his work *The Republic* to illustrate "our nature in its education and want of education" (514a). It is written as a fictional dialogue between Plato's teacher Socrates and Plato's brother Glaucon at the beginning of Book VII (chapter IX in Robin Waterfield's translation) (514a–520a). The Allegory of the Cave is presented after the metaphor of the sun (507b–509c) and the analogy of the divided line (509d–513e). Allegories are summarized in the viewpoint of dialectic at the end of Book VII and VIII (531d–534e). Plato lets Socrates describe a group of people who have lived chained to the wall of a cave all of their lives, facing a blank wall. The people watch shadows projected on the wall by things passing in front of a fire behind them, and begin to ascribe forms to these shadows. According to Plato's Socrates, the shadows are as close as the prisoners get to viewing reality. He then explains how the philosopher is like a prisoner who is freed from the cave and comes to understand that the shadows on the wall do not make up reality at all, as he can perceive the true form of reality rather than the mere shadows seen by the prisoners.
A number of organizations had been following this trend closely - including my own (ITHAKA ... which is the organizational umbrella under which JSTOR and Portico reside). We were taking a longitudinal look at faculty views about the library – and other pertinent scholarly communications issues – and comparing those view with similar survey data from librarians.

One noticeable disconnect in these surveys – as you might imagine – as the perception of the “library as gateway”. Librarians believe it to be hugely important and faculty less so (science faculty much less so than humanities faculty). And students? Even less than that.

Yet, the dollars being spent on access services in libraries – both software and people – were (and continue to be) tremendous. Are those expenditures aligned properly with the expectations of the users, and if they are, then how do we more effectively leverage those investments to reach a broader audience?
A number of organizations had been following this trend closely - including my own (ITHAKA … which is the organizational umbrella under which JSTOR and Portico reside). We were taking a longitudinal look at faculty views about the library – and other pertinent scholarly communications issues – and comparing those view with similar survey data from librarians.

One noticeable disconnect in these surveys – as you might imagine – as the perception of the “library as gateway”. Librarians believe it to be hugely important and faculty less so (science faculty much less so than humanities faculty). And students? Even less than that.

Yet, the dollars being spent on access services in libraries – both software and people – were (and continue to be) tremendous. Are those expenditures aligned properly with the expectations of the users, and if they are, then how do we more effectively leverage those investments to reach a broader audience?
Is the juice worth the squeeze?

How Users Search the Library from a Single Search Box

http://crl.acrl.org/content/early/2012/01/09/crl-321.short?rss=1

Lown, Sierra, Boyer
NCSU Libraries
to be published in C&RL,
March 2013
Where is activity originating in JSTOR?

Linking Partners: link resolvers (e.g. SFX); linking partners (e.g. RePEc, philpapers, etc.); publishers (e.g. AMS)

Academic: library websites; LibGuides; Course management systems; proxy servers

Non-Google search: Bing, Microsoft Academic Search, Baidu, Ask.com

Other: Facebook, Wikipedia, RenRen
What is the early impact on JSTOR usage?

### Web-Scale Discovery Use: 2010 vs. 2011

<table>
<thead>
<tr>
<th></th>
<th>Total Accesses</th>
<th>Searches</th>
<th>Article Views</th>
<th>PDF Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>JSTOR</td>
<td>-2%</td>
<td>+7%</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>Summon</td>
<td>-14%</td>
<td>-23%</td>
<td>+3%</td>
<td>-3%</td>
</tr>
<tr>
<td>Primo</td>
<td>-13%</td>
<td>-12%</td>
<td>-3%</td>
<td>-3%</td>
</tr>
<tr>
<td>EDS</td>
<td>-5%</td>
<td>+28%</td>
<td>-9%</td>
<td>0%</td>
</tr>
<tr>
<td>WorldCat</td>
<td>-9%</td>
<td>+7%</td>
<td>-7%</td>
<td>+1%</td>
</tr>
</tbody>
</table>

Searches: standard deviation is large for EDS (65%); XML searches dominate.
What is the early impact on JSTOR usage?

Figure 1: Content accesses referred from indexed discovery services as a % of overall JSTOR usage
What is the early impact on JSTOR usage?

Figure 2: Content accesses referred from individual indexed discovery services as a % of overall JSTOR usage

This jump in usage sourced from Summon is almost certainly explained by the Summon platform’s introduction of direct linking at this time (rather than routing all traffic through a link resolver). This helps demonstrate some of the complexities of tracking usage based on referrers; this apparent rise in traffic from Summon may in large part simply reflect the removal of an intermediary that previously masked traffic sourced from Summon.
What is the early impact on JSTOR usage?

Figure 3a: US research institution traffic sourcing for content accesses, 2011, Summon example

Summon traffic picking up after direct linking implemented
What is the early impact on JSTOR usage?

Figure 3b: UK research institution traffic sourcing for content accesses, 2011, Primo example

Usual Google/Scholar traffic
What is the early impact on JSTOR usage?

Figure 3c: ARL institution traffic sourcing for content accesses, 2011, EDS example

Highlights problem with accurately capturing origins, as activity is directed via proxy server
What is the early impact on JSTOR usage?

Figure 3d: ARL institution traffic sourcing for content accesses, 2011, WorldCat example

Highlights higher percentage of “Library” traffic ... perhaps driven from WorldCat, but disguised as “Library”?
The highest average content accesses per session referrers are library web pages and other academic institutional websites. Mainstream sites (including Google, Google Scholar, Wikipedia, and more) all have a significantly lower average number of content accesses per session. Indexed discovery services all fall in between, a second tier of referrers behind library and institutional websites. Summon and EDS have a slightly higher rate of average content accesses per session than do Primo or WorldCat. The WorldCat numbers may be explained by the inclusion of worldcat.org in this category; Primo is harder to explain, and may relate to the added complexity of categorizing Primo referrers, indicate that Primo is presenting users with links to materials that they may not have access to, or be an artifact of the fact that Primo is typically integrated into the library website in a way that makes Primo sites more accessible to unauthenticated users.
What is the early impact on JSTOR usage?

Percentage of referrals resulting in turnaways. As expected, sites available to mainstream users result in the highest share of turnaways, with more targeted academic sites that often require authentication before a user is directed to JSTOR resulting in a far lower share of turnaways. As above, Primo and WorldCat demonstrate a slightly different pattern than the other discovery services.
Figure 10: EDS default settings does not show article in JSTOR if it is also available in EBSCO database

EDS only shows link to article in EP database
Issues: Categorization

Figure 9: EDS categorizing the same journal in different ways
Figure 11a: JSTOR metadata is insufficient for Summon metadata schema and relevancy algorithm
Figure 11b: JSTOR metadata not optimized for Primo metadata schema and relevancy algorithm
Issues: Signaling

Figure 12: Does any non-librarian know what JSTOR Arts & Sciences I Archive Collection is?
Next Steps

- Expand usage analysis to more institutions using each of the discovery services, as well as compare to institutions not using a discovery service;
- Work with discovery providers to more accurately and consistently provide origin designators;
- Gain deeper understanding of how library admin modules work in the various services and provide input to participating libraries and publishers;
- Improve scope of metadata being delivered to discovery service providers; optimize metadata delivered to provider-specific metadata schema and relevancy algorithms;
- Experiment with providing more than metadata to providers and measure change in usage; if measurably different, consider implementing with preferred partners;
- Reward good behavior amongst the discovery providers; expose bad behavior.