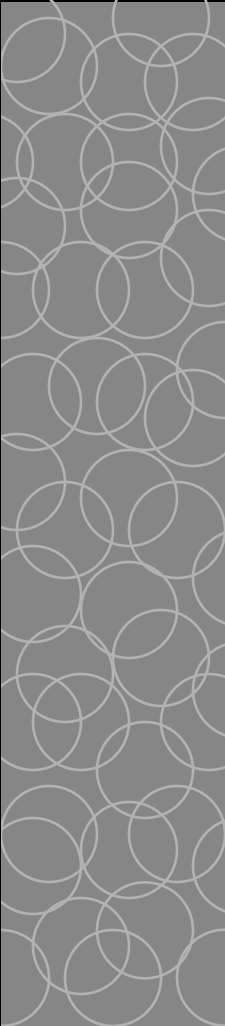




Privacy and Library Systems

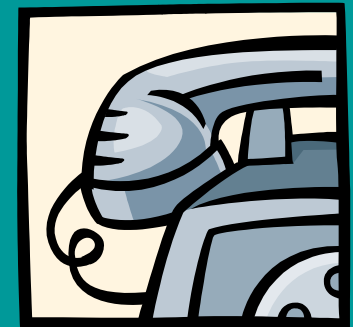
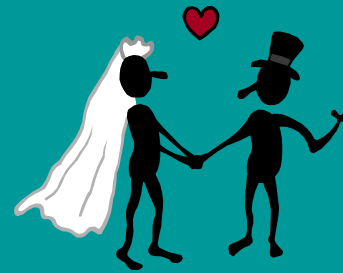
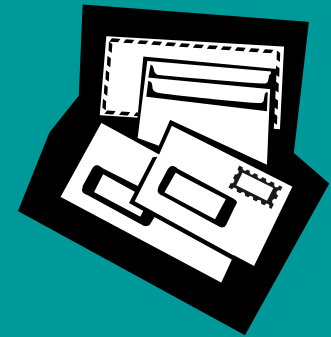


*Karen Coyle
for InfoPeople
November, 2004*

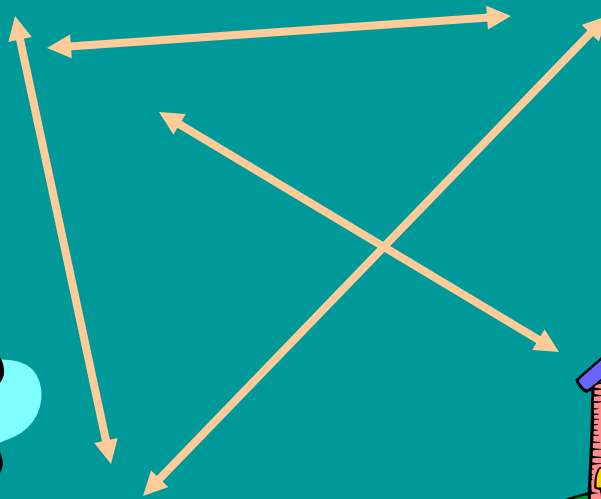
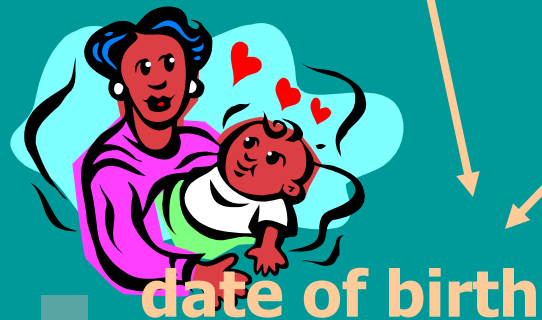
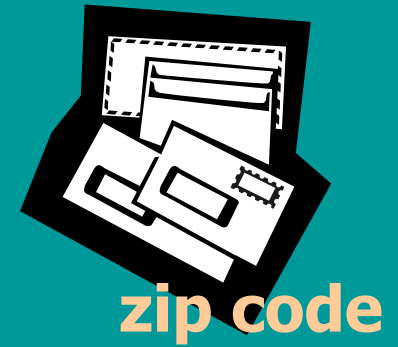
Privacy Today



111-22-3333

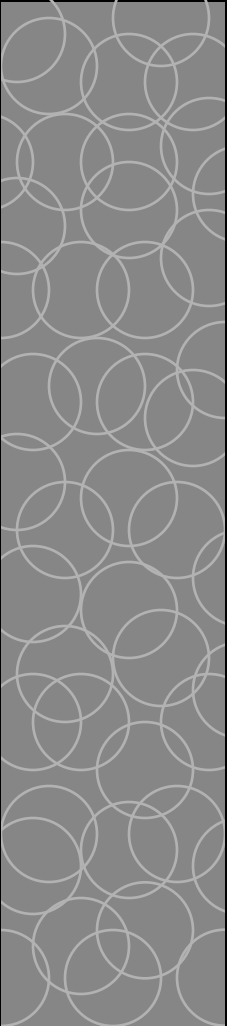


"Data Mining"



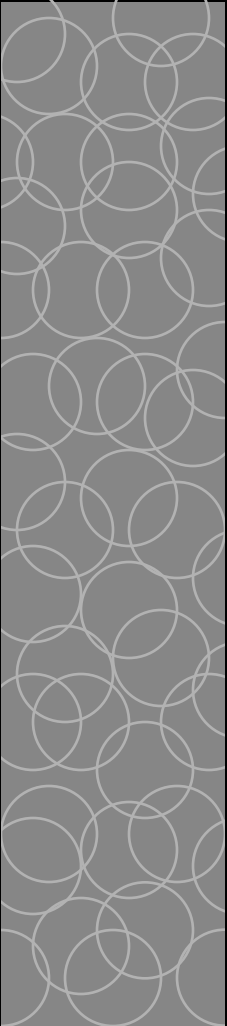


Privacy Today – in Libraries

- 
- ◆ *Computer systems create records where none existed before*
 - ◆ *Library services go beyond library walls*
 - ◆ *Library network is part of a larger networked world*

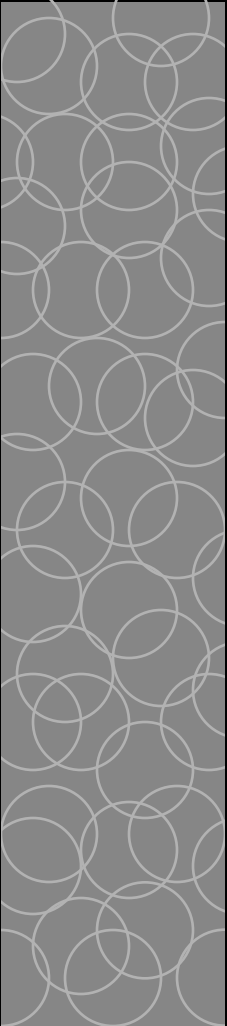


Privacy Today – in Libraries

- 
- ◆ *Computer systems create records where none existed before*
 - ◆ *Library services go beyond library walls*
 - ◆ *Library network is part of a larger networked world*
 - ◆ *Any library data is in the context of the “data soup”*



What libraries can do....



*Be responsible keepers of
information about our patrons
and library use*

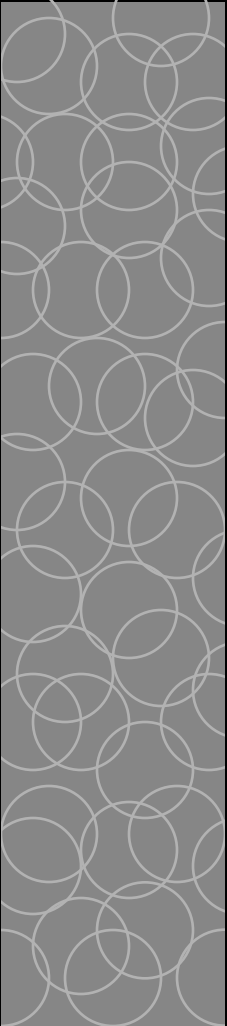


Be honest and informative





Service and Privacy



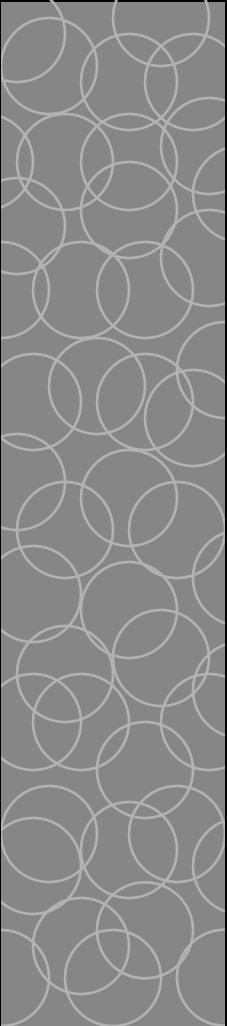
*More service can require more
user information*



*“Personalization” is ...
personal*



Security and Privacy



*An insecure system cannot
protect the privacy of users*

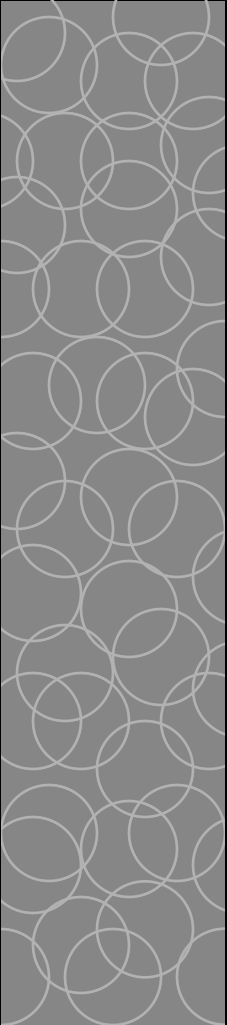
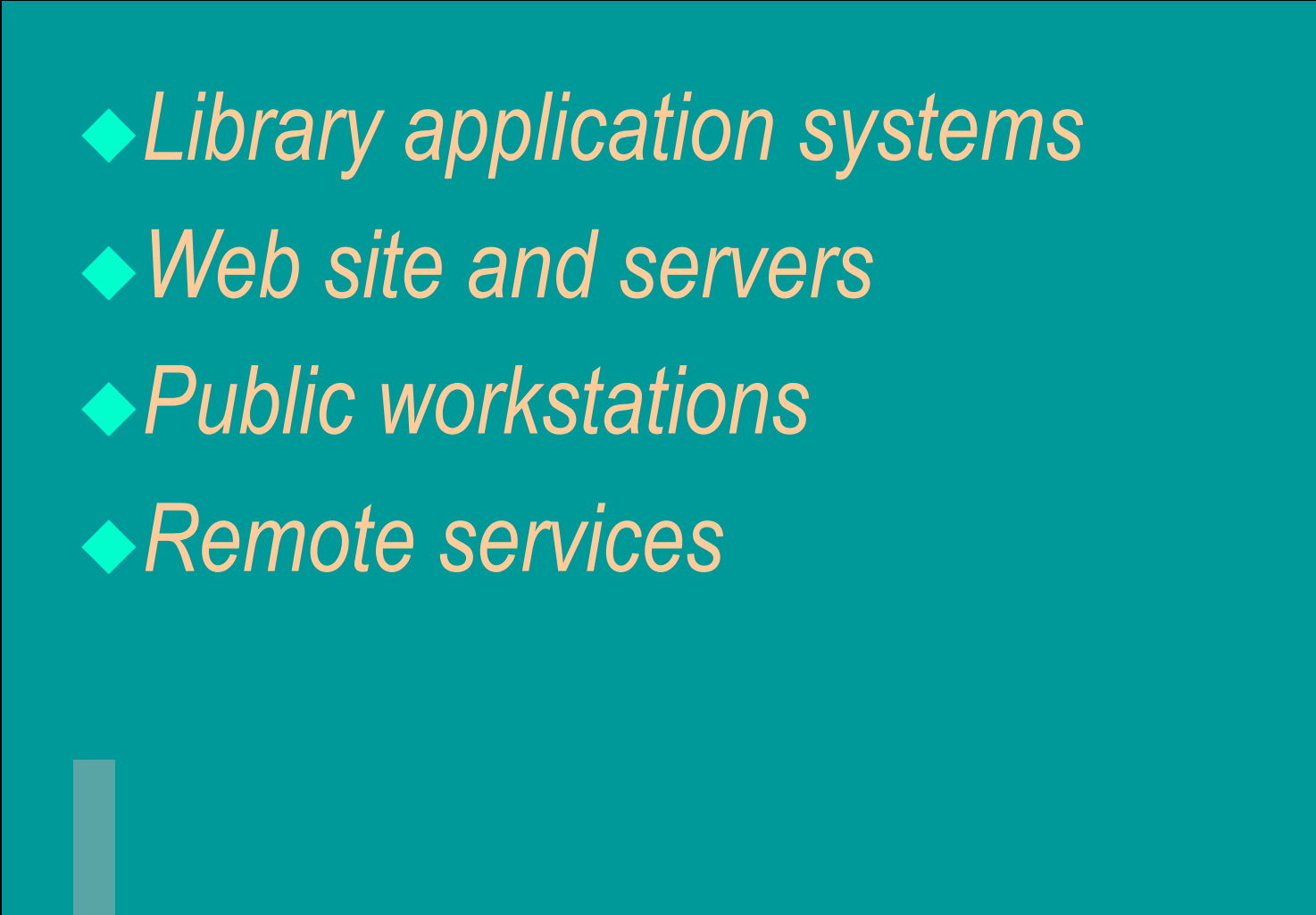


*Desire for security can lead to
monitoring and logging of use*



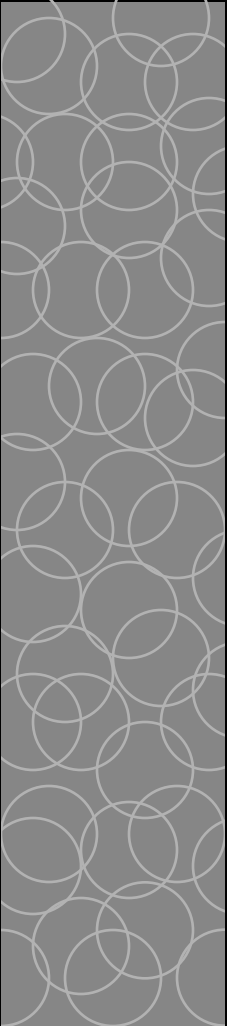


Systems & Privacy

- 
- ◆ *Library application systems*
 - ◆ *Web site and servers*
 - ◆ *Public workstations*
 - ◆ *Remote services*
- 



Rules for data collection

- 
- ◆ *Only gather data that is absolutely necessary to perform the function*
 - ◆ *Do not keep data longer than is needed to perform the function*
 - ◆ *Limit access to those persons who must work with the data*



Library Application Systems



Circulation and Borrower Records

- ◆ *Patron database*
- ◆ *Circulation records*
- ◆ *Fines and payments*
- ◆ *Document delivery*
- ◆ *ILL requests and fulfillments*
- ◆ *Electronic reserves*
- ◆ *OPAC logs*



Library Application Systems

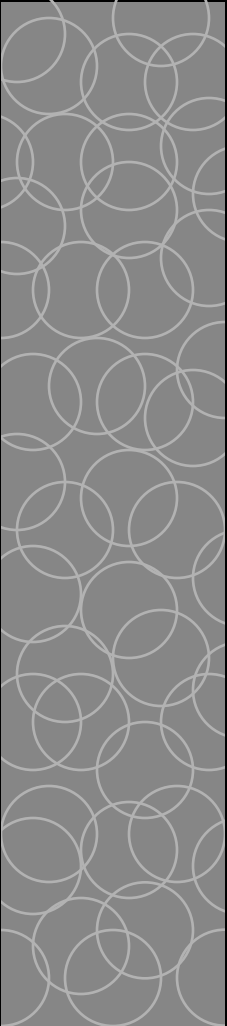



Personalized Services

- ◆ *Saved searches or bibliographies*
- ◆ *Automated search profiles (SDI)*
- ◆ *Virtual reference*
- ◆ *“my library” personal view*
- ◆ *email services*

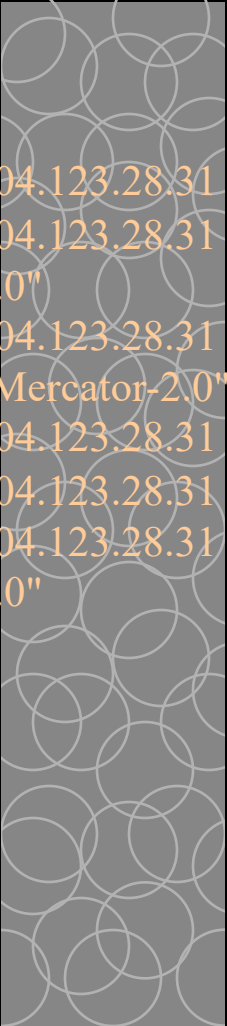


Library Web Site

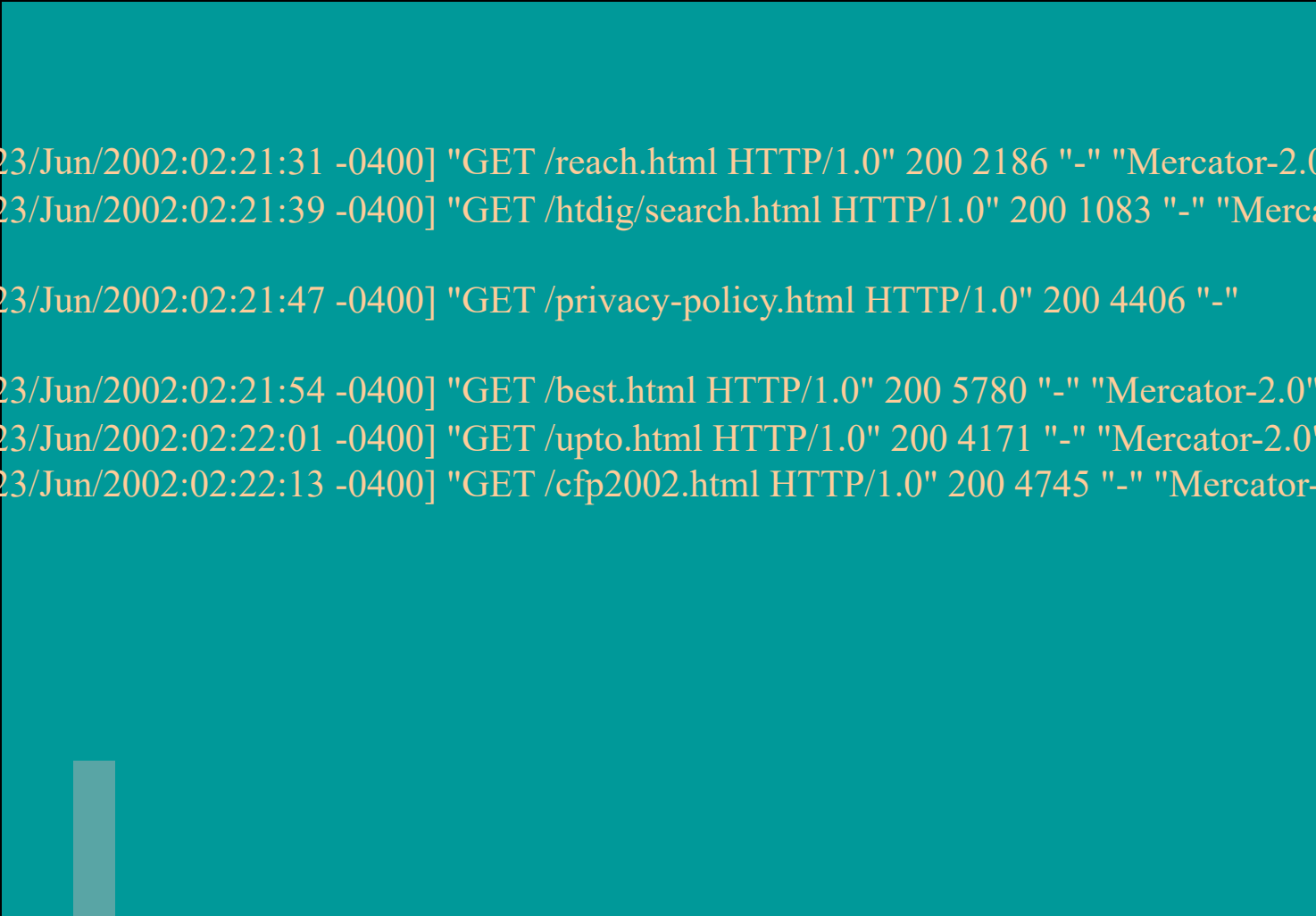
- 
- ◆ *Web server logs*
 - ◆ *Proxy server logs*
 - ◆ *Cookies*
 - ◆ *Forms and email*
- 



Sample web log



```
204.123.28.31 - - [23/Jun/2002:02:21:31 -0400] "GET /reach.html HTTP/1.0" 200 2186 "-" "Mercator-2.0"
204.123.28.31 - - [23/Jun/2002:02:21:39 -0400] "GET /htdig/search.html HTTP/1.0" 200 1083 "-" "Mercato
2.0"
204.123.28.31 - - [23/Jun/2002:02:21:47 -0400] "GET /privacy-policy.html HTTP/1.0" 200 4406 "-"
"Mercator-2.0"
204.123.28.31 - - [23/Jun/2002:02:21:54 -0400] "GET /best.html HTTP/1.0" 200 5780 "-" "Mercator-2.0"
204.123.28.31 - - [23/Jun/2002:02:22:01 -0400] "GET /upto.html HTTP/1.0" 200 4171 "-" "Mercator-2.0"
204.123.28.31 - - [23/Jun/2002:02:22:13 -0400] "GET /cfp2002.html HTTP/1.0" 200 4745 "-" "Mercator-
2.0"
```



Sample web log

```
204.123.28.31 - - [23/Jun/2002:02:21:23 -0400] "GET /topiclist.html HTTP/1.0" 200 17602 "-" "Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""
```



Internet address of user: **204.123.28.31**

Date and time: **[23/Jun/2002:02:21:23 -0400]**

Requested document: **"GET /topiclist.html "**

Browser: **"Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""**

Sample web log

```
204.123.28.31 - - [23/Jun/2002:02:21:23 -0400] "GET /topiclist.html HTTP/1.0" 200 17602 "-" "Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)"
```



Internet address of user: **204.123.28.31**

Date and time: **[23/Jun/2002:02:21:23 -0400]**

Requested document: **"GET /topiclist.html "**

Browser: **"Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)"**

Sample web log

```
204.123.28.31 - - [23/Jun/2002:02:21:23 -0400] "GET /topiclist.html HTTP/1.0" 200 17602 "-" "Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""
```



Internet address of user: **204.123.28.31**

Date and time: **[23/Jun/2002:02:21:23 -0400]**

Requested document: **"GET /topiclist.html "**

Browser: **"Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""**

Sample web log

```
204.123.28.31 - - [23/Jun/2002:02:21:23 -0400] "GET /topiclist.html HTTP/1.0" 200 17602 "-" "Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""
```



Internet address of user: **204.123.28.31**

Date and time: **[23/Jun/2002:02:21:23 -0400]**

Requested document: **"GET /topiclist.html "**

Browser: **"Mozilla/3.01-C-MACOS8 (Macintosh; I; PPC)""**





Sample Web Log

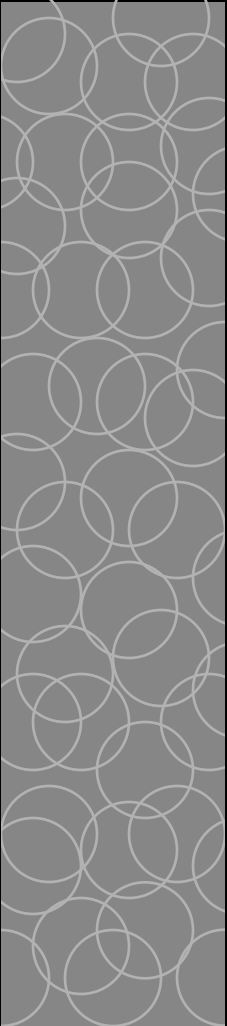
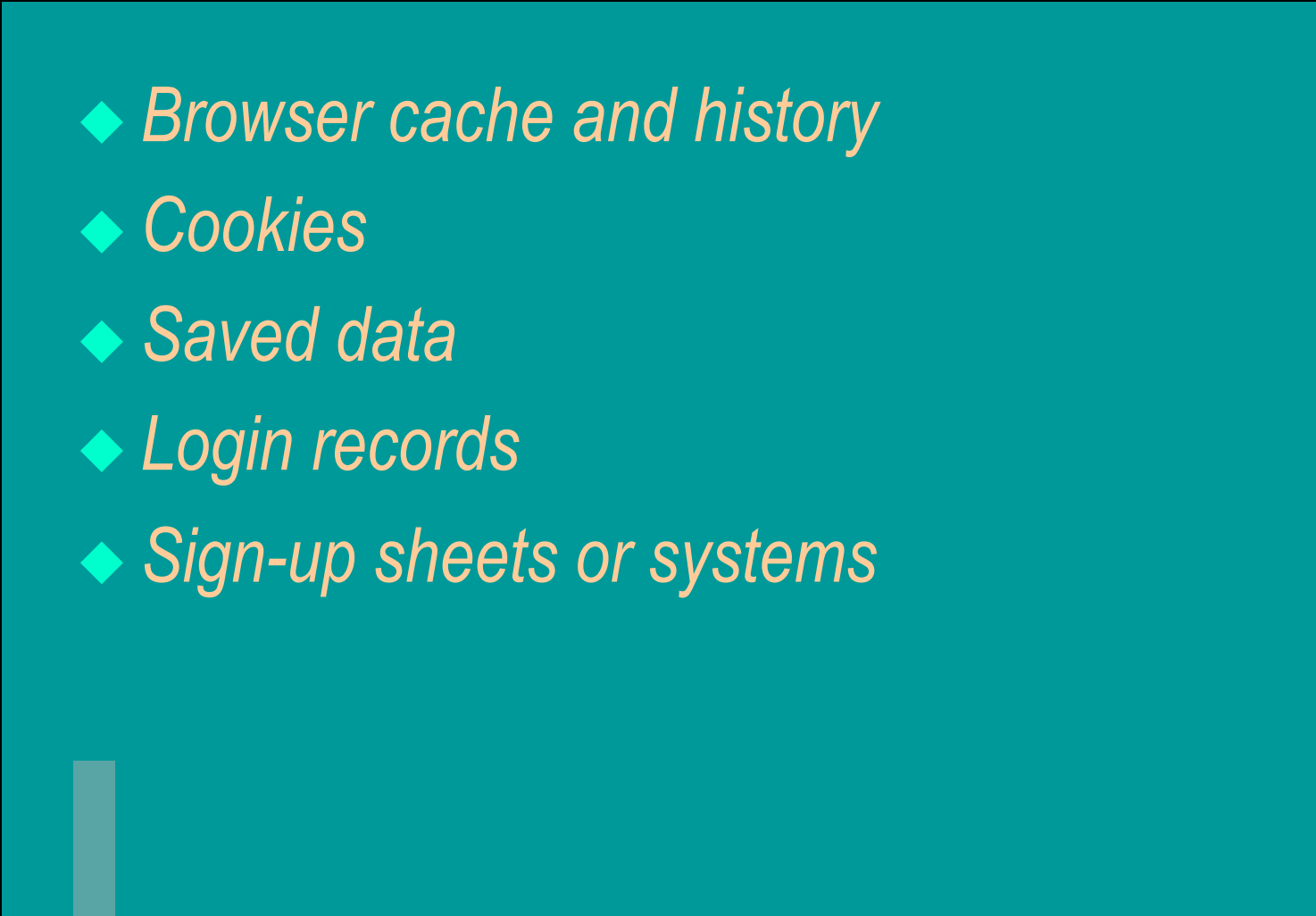


Can include searches

```
"http:www.google.com/search?hl=en&ie=UTF-8  
&oe=UTF8&q=library+not+safe  
+place"
```



Public Workstations

- 
- ◆ *Browser cache and history*
 - ◆ *Cookies*
 - ◆ *Saved data*
 - ◆ *Login records*
 - ◆ *Sign-up sheets or systems*
- 



Remote Services



Databases and licensed services

- ◆ *Logons (in library or from home)*
- ◆ *Personalization*
- ◆ *Statistics*



Remote Services



Internet use

The Internet is outside
of the library



User education





Privacy audit materials



<http://www.kcoyle.net/infopeople/>



Library Systems Privacy Inventory

Circulation and Borrower Records

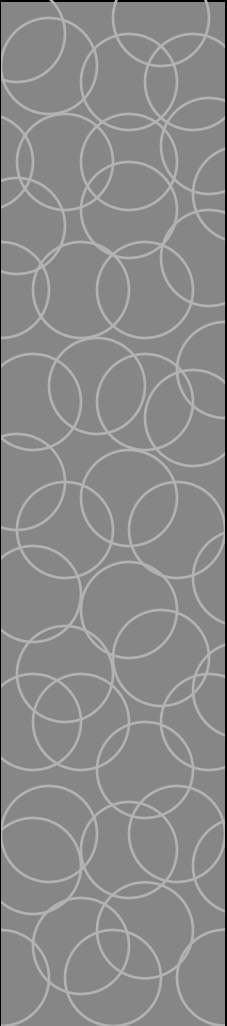
Data	What data is recorded?	Where is it located?	Who has access?	How long is data kept?	Recommendations
Patron database					
Circulation transactions					
Fines and other billing					
Document delivery					
ILL transactions					
Electronic reserves					

OPAC and System Services

Data	What data is recorded?	Where is it located?	Who has access?	How long is data kept?	Recommendations
Search histories saved beyond a session					
Saved searches and					



For each:

- 
- ◆ *What data exists?*
 - ◆ *Where is it located?*
 - ◆ *Who has access?*
 - ◆ *How long is the data kept?*



For each:



- ◆ *What data exists?*

What is gathered and what does reveal?

- ◆ *Where is it located?*

Is it within the library's control?

- ◆ *Who has access?*

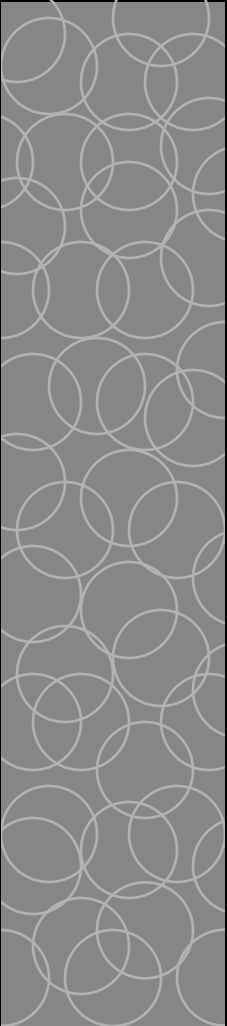
Who *must* have access to do their job?

- ◆ *How long is the data kept?*

How long is this data *really* needed?



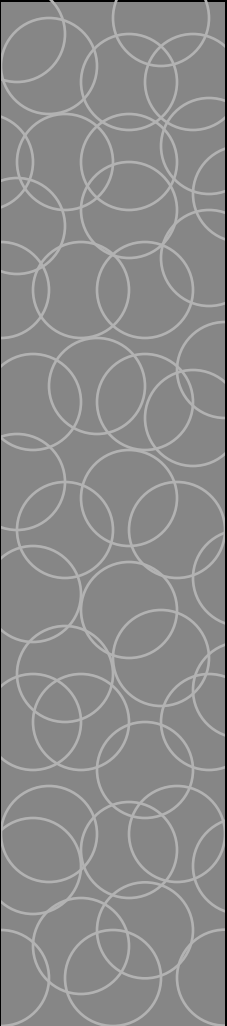
Example: Saved bibliography



<i>What?</i>	<i>Who?</i>	<i>How long?</i>
<i>List of record IDs + Patron ID + Date created. Saved on server</i>	<i>Patron has access to their own; system admin has access to all; no other staff access</i>	<i>Purged at the end of each academic year; patrons can delete</i>

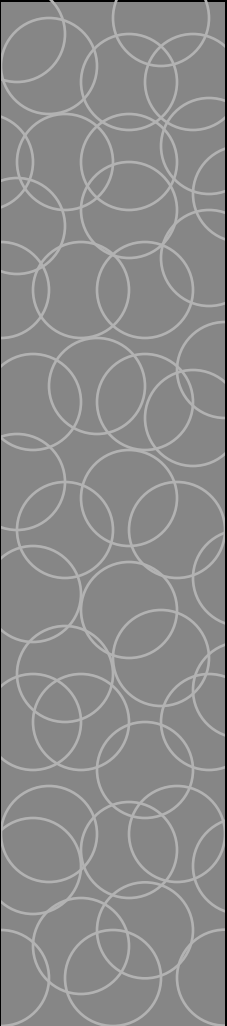


Example: Saved bibliography

- 
- ◆ *Remind users to delete saved bibliographies they are no longer using*
 - ◆ *Remove bibliographies for user accounts that are no longer active*
 - ◆ *Inform users that although the library wishes to provide a secure environment, it cannot guarantee the privacy of these files.*



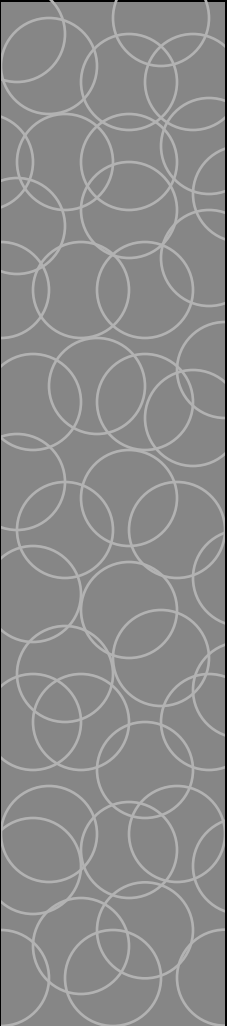
Example: Web logs



<i>What?</i>	<i>Who?</i>	<i>How long?</i>
<i>Each request sent to library web server; includes: IP address of sender; item requested; browser...</i>	<i>Only systems staff have access</i>	<i>Stay on server until log file fills up; about 6 weeks</i>

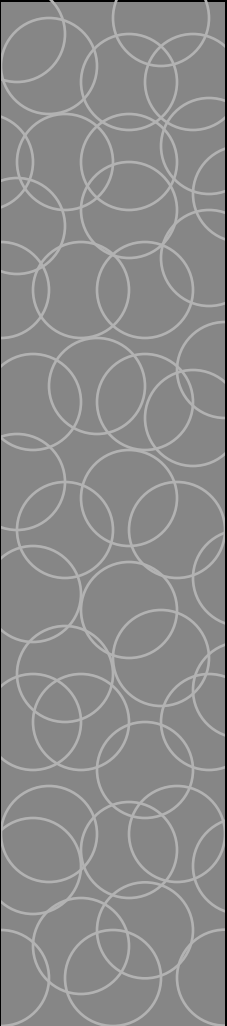



Example: Web logs

- 
- ◆ *If using web logs for statistics, gather statistics immediately and then purge the logs*
 - ◆ *If logs must be kept, “anonymize” the user’s identity by removing all or part of the user’s incoming Internet address*



Summary

- 
- ◆ *Identify system areas where data is stored*
 - ◆ *Determine how long it is needed, and who should have access*
 - ◆ *Educate staff and educate users*
- 



Thank You.



Karen Coyle

<http://www.kcoyle.net>

