Content Ratings, and Other Third-Party Value-Added Information: Defining an Enabling Platform

Martin Röscheisen
`rmr@cs.stanford.edu`

Terry Winograd
`winograd@cs.stanford.edu`

Andreas Paepcke
`paepcke@cs.stanford.edu`

Stanford Integrated Digital Library Project
Computer Science Department
Stanford University
Stanford, CA 94305, U.S.A.

> *Scenario:* Imagine you want to know what your colleagues in the interest group `DLissues` have found worth seeing lately. In your browser, you select "Tour annotation set `DLissues`", with the filter set to "annotations newly created since yesterday". You get a report containing pointers to annotated locations in various documents; you inspect some of these links with a comment previewer. Sara evidently appreciated a paper on security in the proceedings of a conference last year--she gave it the highest ranking on her personal scale. You click on the link and jump to the annotated section in the paper. You scan it up and down and wonder whether the security research group you know at another university has any opinion on this paper. You turn on their annotation set `SecurityPapers`, which you can access free of charge since your school has a site licensing agreement. You see that they have made a "trailmarker annotation" to the top of the paper. You inspect the annotation icon with the previewer: it says that the paper you are viewing is really subsumed now by the one at a more recent conference which the trail marker points to. With another click you jump to this more recent paper, which turns out to be written even more clearly. You go back to reply to Sara's original comment and include a pointer to the `SecurityPapers` set.

As part of the Stanford Integrated Digital Library Project, we have designed and prototyped an architecture that supports such scenarios by defining an enabling platform for various kinds of third-party value-added information on top of the existing World-Wide Web infrastructure.

A research prototype implementation ("ComMentor") was completed in late 1994 and has been refined and evaluated since then. The protocols have been published as a spin-off from our digital library project to the World-Wide Web community; the code has been made available to other developers to enable possibilities for wider adoption and for the development of related standards.

A rich variety of usages can be readily realized using a generalized mechanism for shared "annotations," which underlies our architecture. These usages include shared comments, collaborative filtering, seals of approval, guided tours, usage indicators, co-presence, and value-added trails.

Third-Party Value-Added Information and the Need for an Enabling Platform

Figure 1 shows a simplified depiction of what browsing on today's World-Wide Web is like: the user's view is limited at any time by whatever information and pointers a content provider makes available.

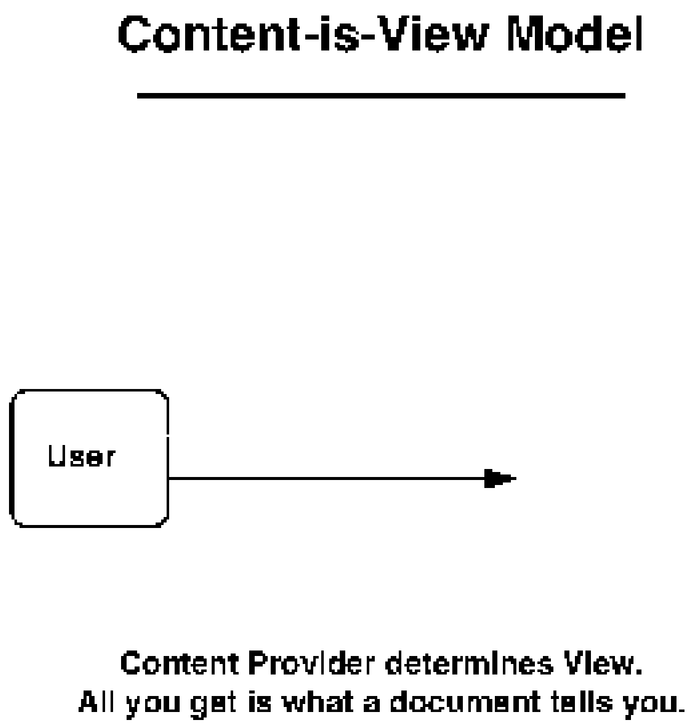


Figure 1

This contrasts to the way in which many kinds of information are used to add value. For example, Consumer Report evaluations about product information, or a professor's comments about a conference paper are generally authored, published, and controlled independently of the underlying content. When users look at the product information, they might want to have a pointer to the corresponding Consumer Report--even if a product manufacturer does not necessarily appreciate a review. Note also that while there might be access control and charging for the product information, the value-added super-structures will in general be access controlled and charged independently.

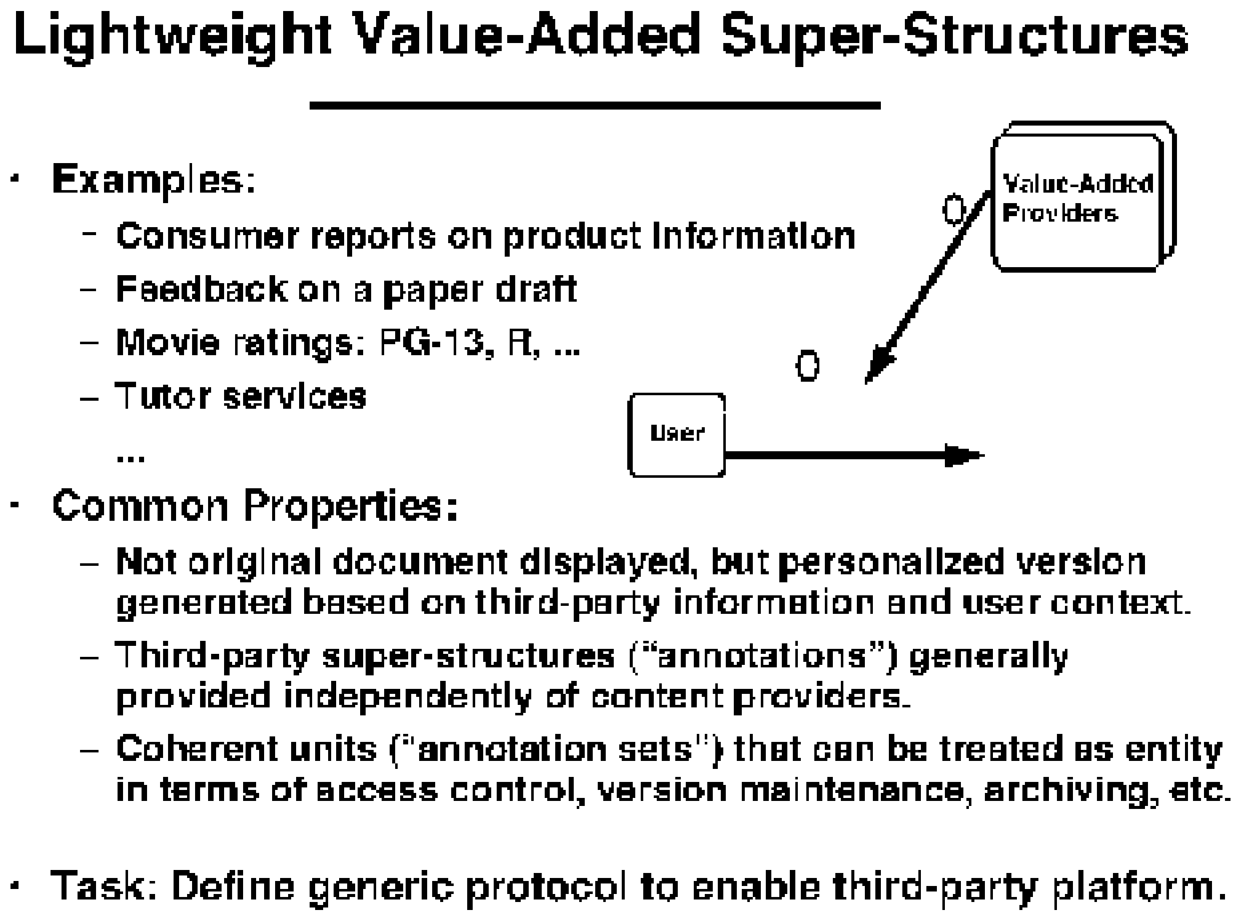


Figure 2

Usages of third-party annotations are not limited to personal annotations. When properly supported by the underlying infrastructure, they also help in the context of collaborative work groups: annotations can indicate which participants of a group have seen a document; they can afford structured discussion about paper drafts and collaborative filtering.

Annotations that include hyperlinks can be used to construct guided tours through a document space, or trails in the form envisioned by Vannevar Bush can be constructed (see also [VB]). Since annotations conceptually reside on pages, such "landmarks" naturally implement a generalized notion of a "hotlist", which is shared among arbitrary groups of people and where visibility/access is controlled on a per-section level. Annotations can also be used by participants to indicate their 'presence' at a document. But one use of annotations has most recently risen to particular prominence: seals of approval (SOAPs) and their use for content ratings.

Content ratings have been debated intensively as part of the excitement around the Communications Decency

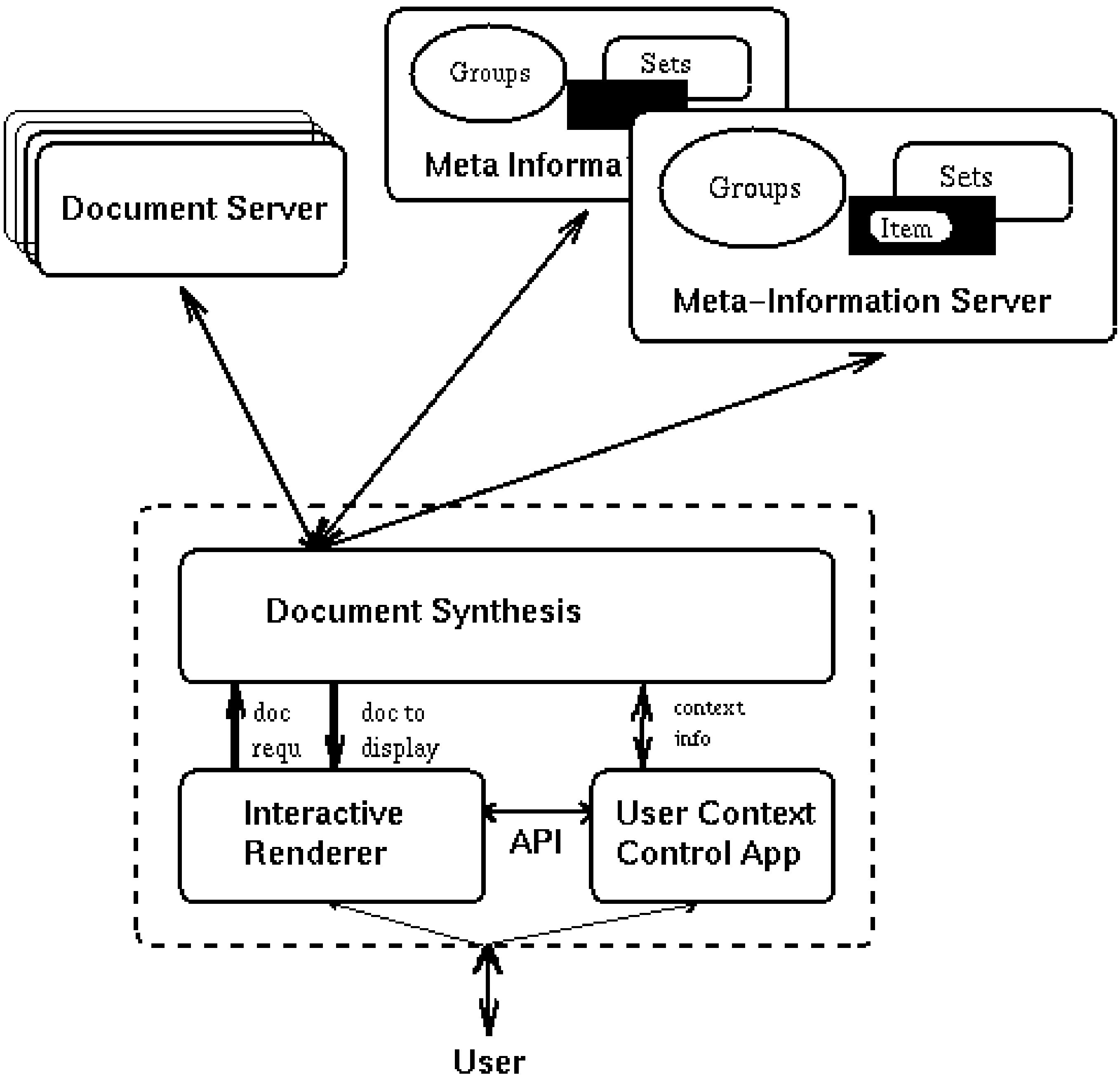
This flexibility becomes even more of an issue when considering that a given document may be rated for very different purposes. The Playboy(TM) Web site might be rated "guidance advised" by a committee of parents; it might also be rated "picture quality good" by an organization of professional photographers, or it might be rated "slow site" by someone who rates sites according to their latency. Content rating annotations are intrinsically relative to perspective, value system, and intended use.

These properties of third-party value-added information make it in general undesirable, if not infeasible, to have content providers also provide the associated third-party information, which was prepared differently in terms of authority, intention, access control, burden of resource usage etc. Once we acknowledge the specific nature of third-party value-added information, it becomes clear that *its independence needs to be reflected architecturally*.

The ComMentor architecture enables independent third parties to provide value-added information ("annotations"), and it provides a generic mechanism for users to choose which kind of such information ("annotation sets") will be "superimposed" when viewing documents.

Outline of the Architecture

The basic architecture is shown in Figure 3. Users interact with a "context-control application" in which they select the third-party source and the type of information ("annotation set") they want to see at a given point in time. For example, while browsing for information on personal computers, someone might choose to turn on the `PCratings` annotation set of a well-known computer designer to get hints about some of the salient issues of the various PCs.



Document Server
Groups
Sets
Groups
Sets
Item
Meta-Information Server
Document Synthesis
doc
requ
doc to
display
context
info
Interactive
Renderer
API
User Context
Control App
User

certain access control groups. Annotations are organized into what we call "annotation sets". These are usually like topics, or like threads in newsgroup readers, and they organize annotations in much the same way as directories do with files.

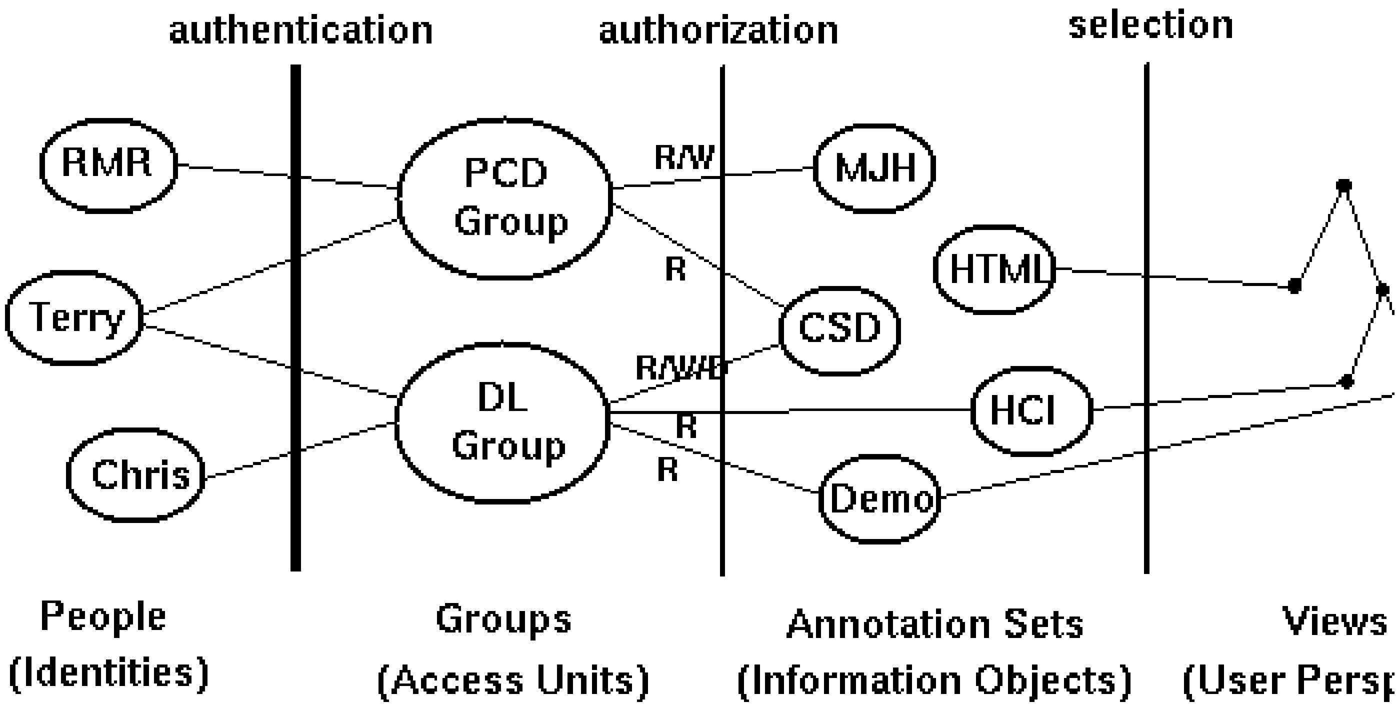


Figure 4

For example, Terry and Chris are members of the `DL Group`. As such they may add or modify annotations in the `CSD set`. But they can only view the annotations in the `Demo set`.

Examples from the ComMentor Prototype

In this section, we will give some examples of how annotations are used in the ComMentor prototype. For a more complete description, see the Technical Report.

Annotations are of different types. There are "comment annotations" for basic commenting, "tour annotations" for guided tours, "SOAP annotations" for content ratings, and others. All annotations contain information such as who authored them when, and which annotation set they belong to. Each of the annotation types defines a number of additional attributes. For example, a tour annotation would have in addition the location to which it points. The type also determines the default client behavior once an annotation is selected. For example, clicking on a "tour annotation" will in general lead to the next tour stop, while clicking on a "comment annotation" will show the full text view of the annotation.

The corresponding meta-information description for an annotation is automatically generated and stored on the annotation server whenever a user uses the "Create Annotation" dialogue box of the browser (Figure 5).

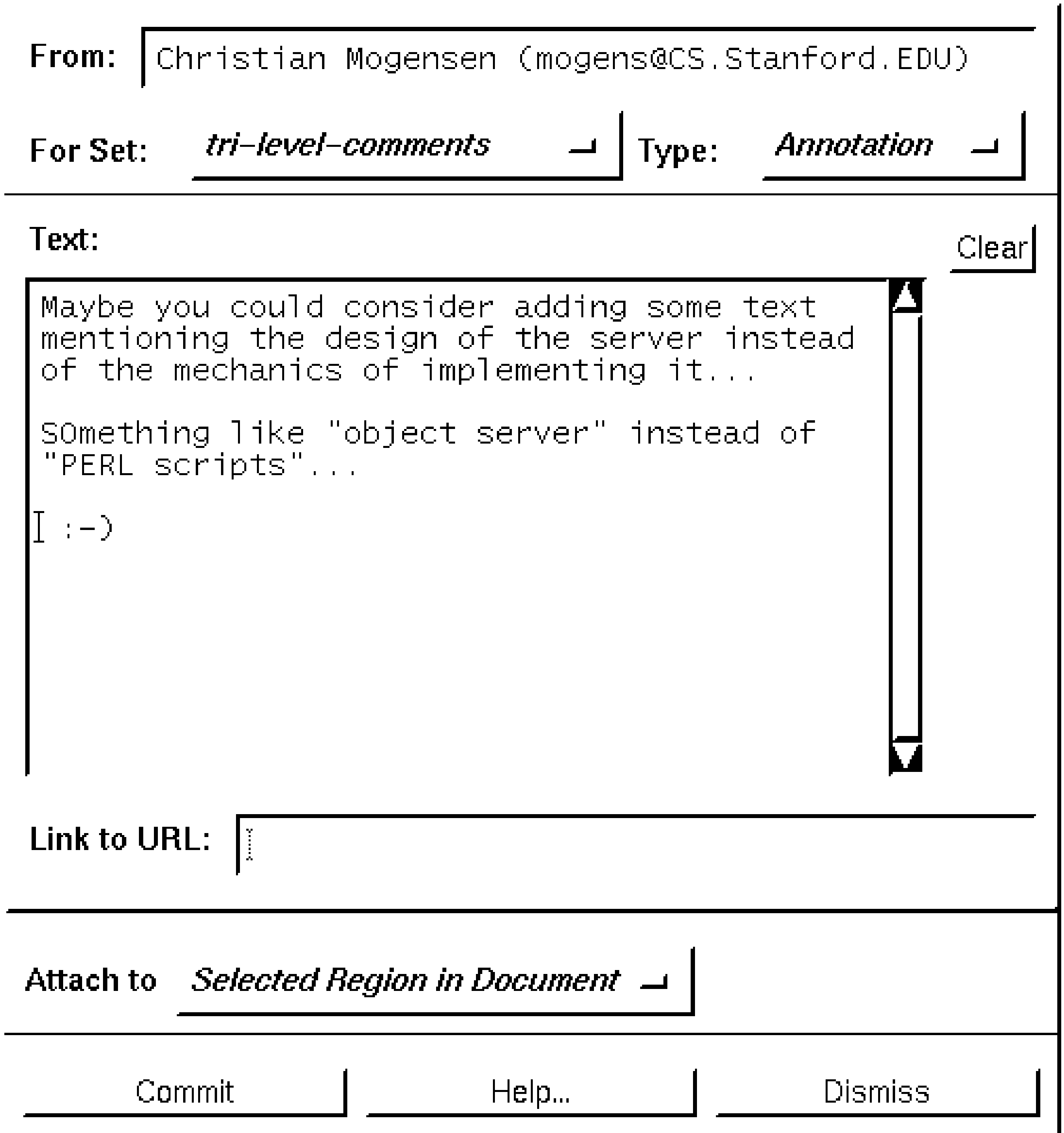


Figure 5

Comment Annotations

Comment annotations are indicated in the interface as tiny icons containing the faces of the author, or (in an alternate viewing mode) an icon of the group to which it was written. Such images are active anchors in that users can click on them to view the comment. We have implemented a previewing mechanism, a yellow

Conclusions

Conclusions

There is much to be gained from considering the nature of a document collection before the authoring process is begun. If the envisioned document collection can be characterized as having recurrent organizational structure which is likely to evolve, then the authoring process for that collection can be improved by choosing a document representation and accompanying set of authoring tools that allow for a semantic, organizational, and presentation distinction to be made.

Michelle Q Wang Baldonado
Fri Dec 2 16:38:51 PST 1994

Figure 6

The picture above shows some comments inlined into the base text (with the highlighted regions shown). Figure 7 shows the previewer being used to inspect an annotation (yellow PostIt viewer).

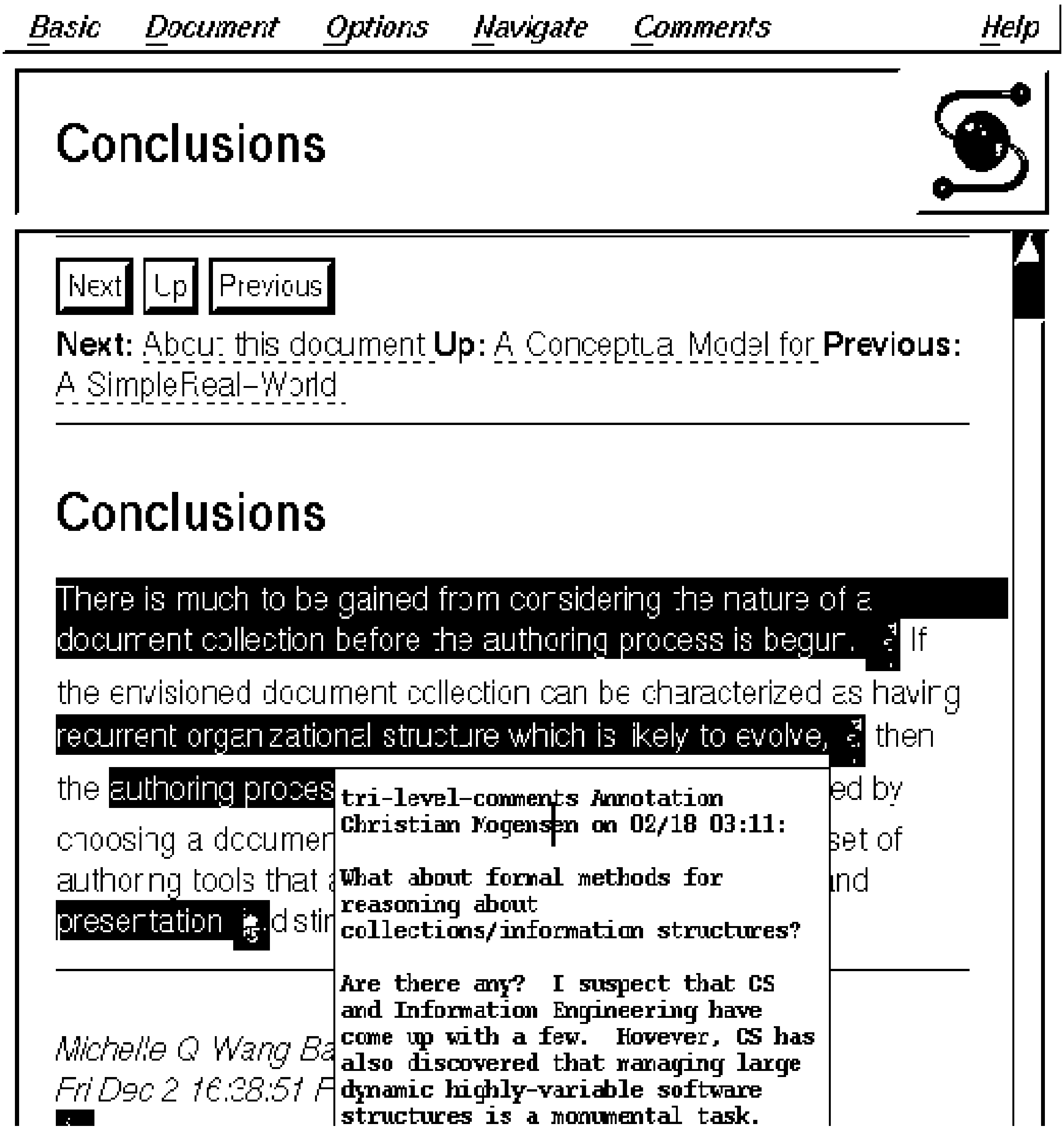


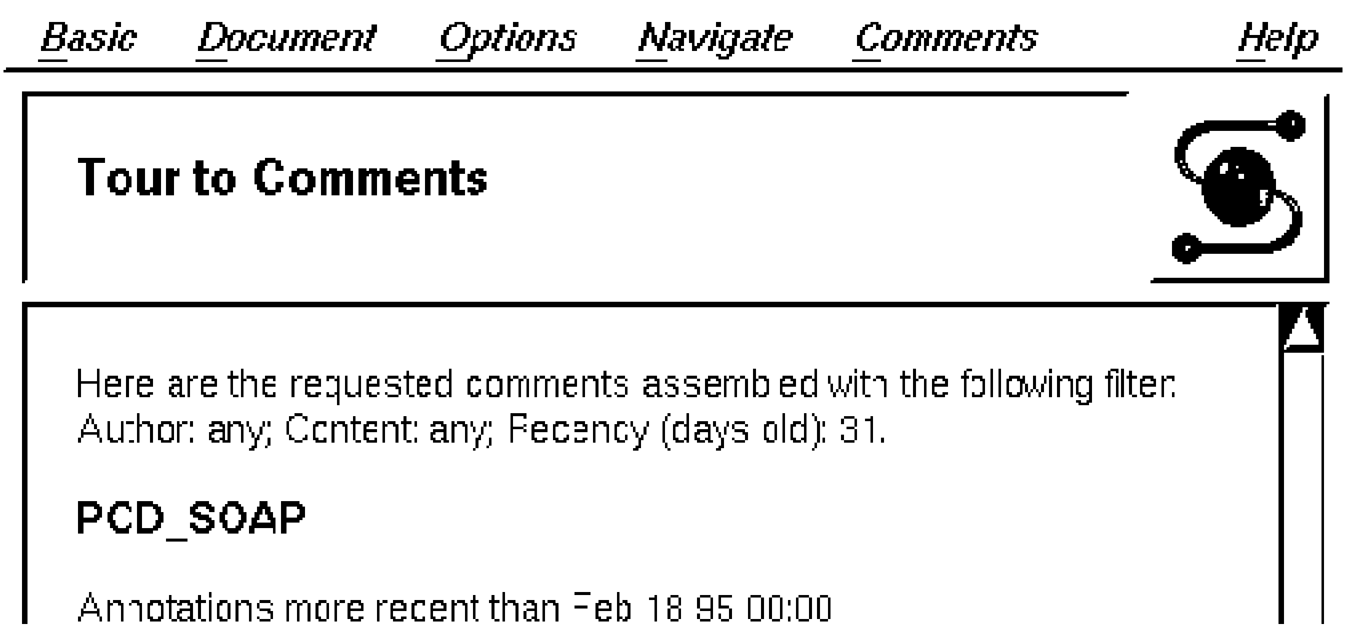
Figure 7

One typical problem with annotations is that they "get lost": The only way to see them is to know the page to which they are attached. We avoid this problem with a general query capability on our annotation servers: users can query for annotations by various criteria. For example, they can ask for all annotations which have been created since yesterday and were written by Andreas. Such a query result shows up in the browser as a "hotlist"-like page where each list item is a link to an annotation.

Seals of Approval (SOAPs) for Content Rating

Content rating is done by writing annotations to pages with respect to specially designated sets whose access control properties are set such that they are readable by whoever the audience is (often: everyone, that is, public), and writable only by whoever belongs to the issuing authority. For example, the French Academy might want to reward especially elegant use of the French language. To that end, it could create a set `LeVraiFrancais` whose access permissions are set such that only academy fellows have write access, but everyone in the public has read access. Anyone on the Web could then turn on this set, and gain insight as to the extent to which a particular document is written in proper French.

There are two basic usages of rating sets: First, using the ability to query the annotation server for a list of pointers to annotations, we help people find what they are looking for. In the following figure, we have queried the `PCD_SOAP` for a list of ratings, among which we can preview the more detailed ratings and then jump to a location of our choice.



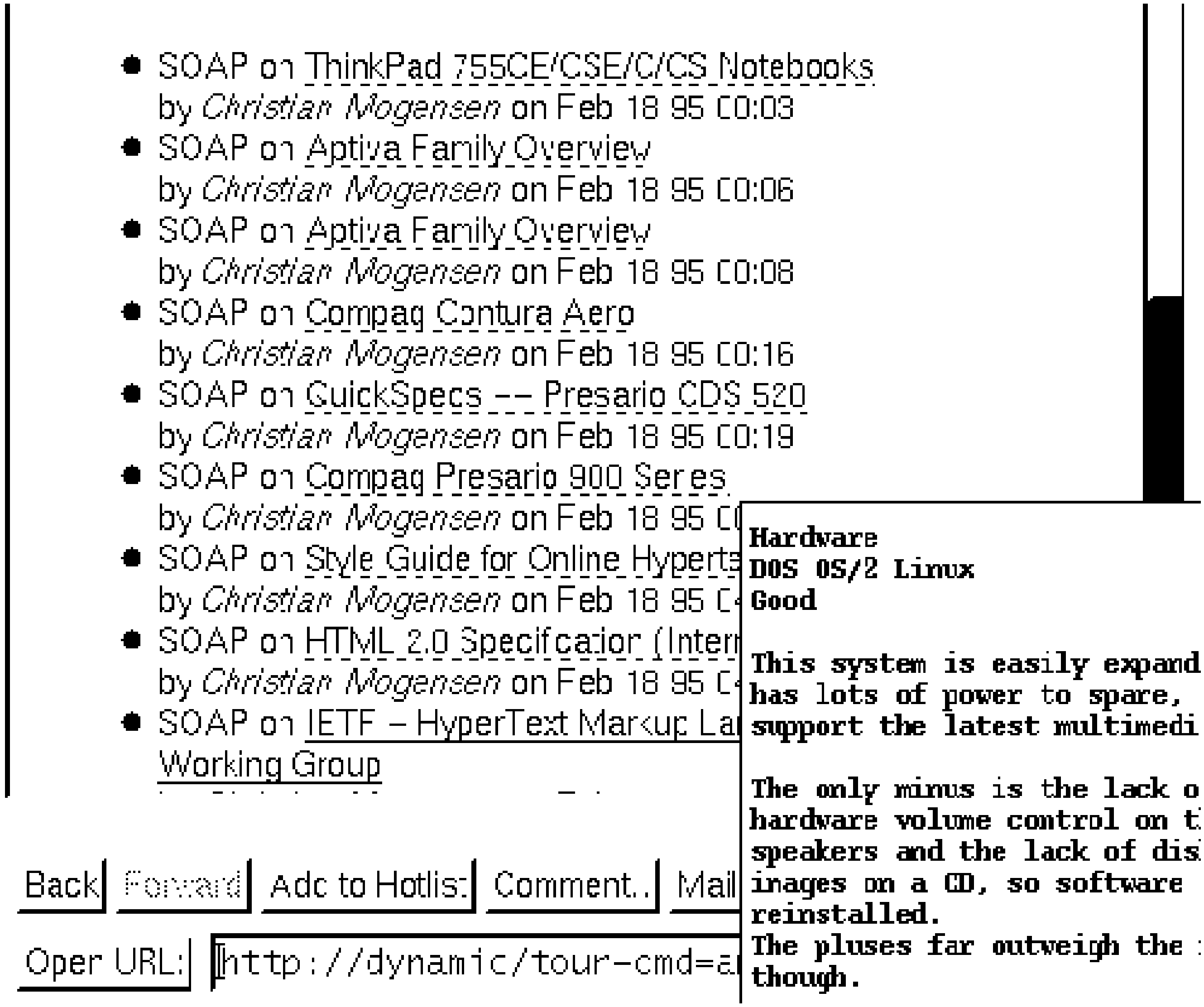
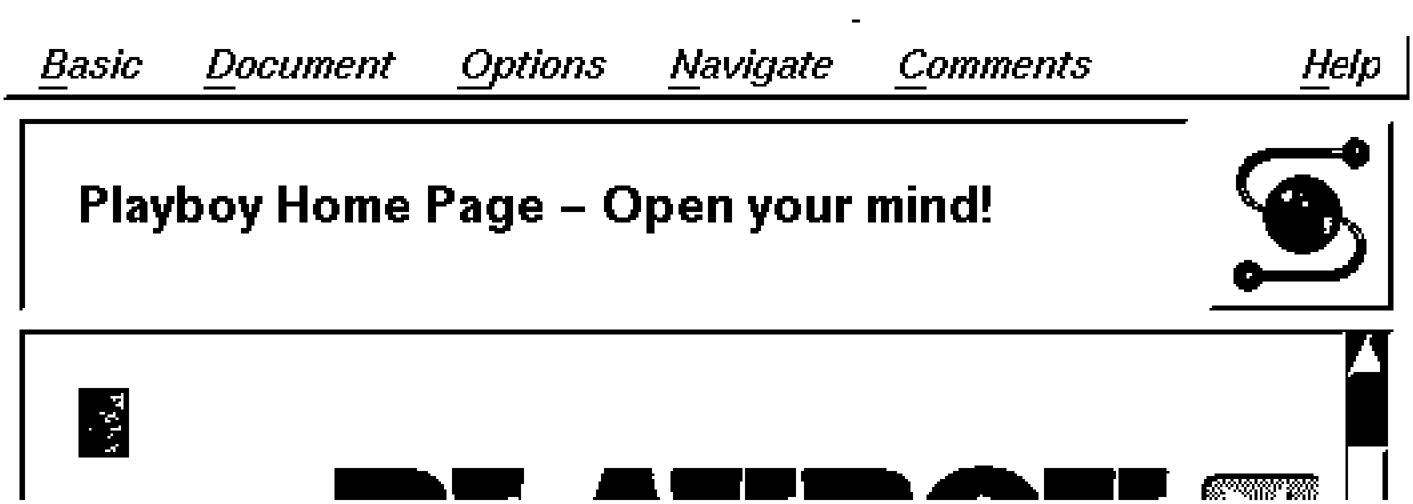


Figure 8

The other main usage of SOAPs is to give people information about something once they happen to run into it. This includes the typical parental guide application (see figure below). Note that here we pop up a notification window; a more useful action might be in this context not to show the page at all.



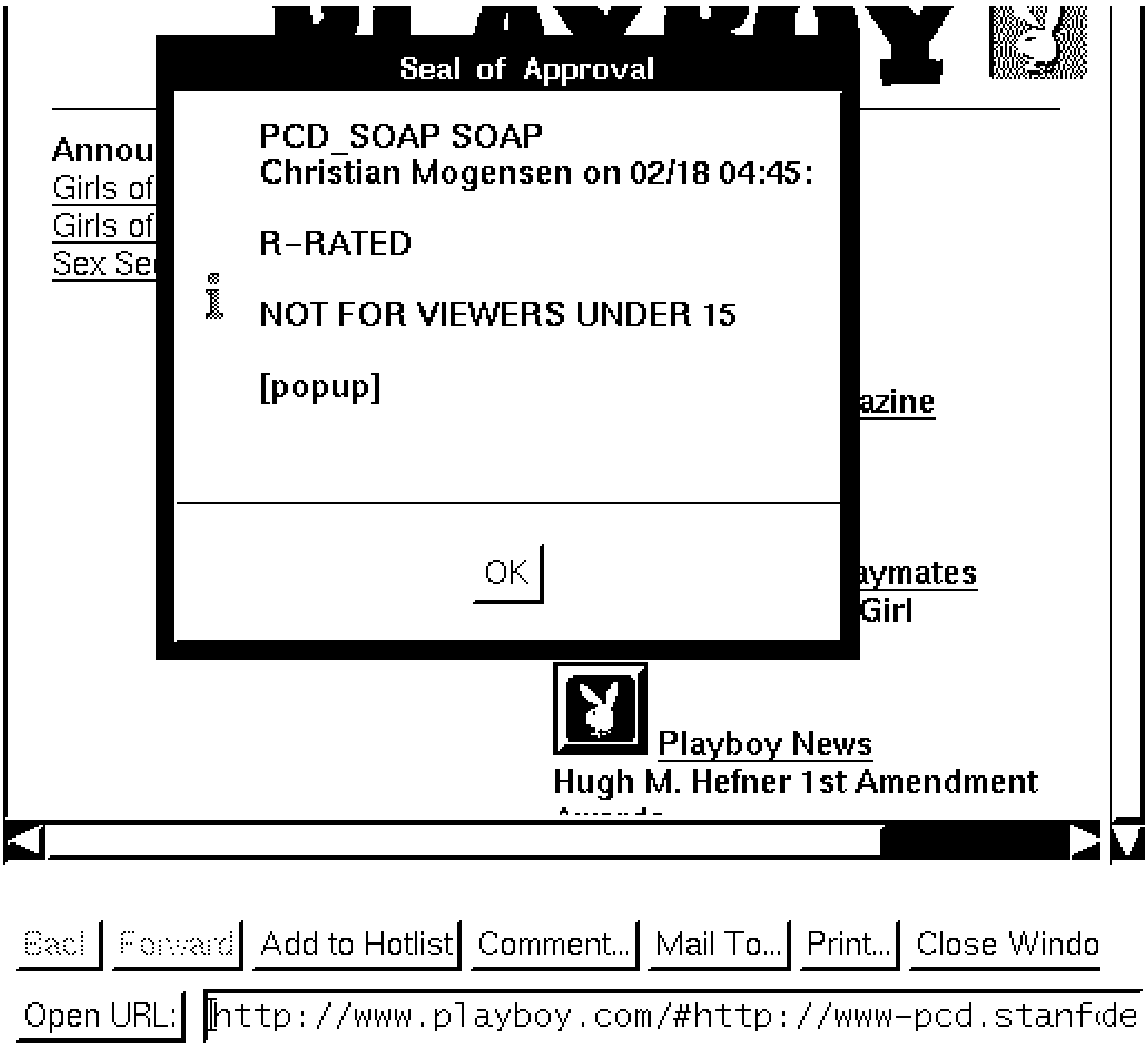


Figure 9

We have implemented a basic set of visibility controls for Seals of Approval which use annotation information to perform extra actions on the client side. These include not showing the underlying document at all (which would be useful in the parental guide case), or popping up a window with a warning message.

Note that the general SOAP structure accommodates any rating scheme: the rating system itself is described as part of the meta-information describing a rating set. For example, a simple annotation set might be created to contain rating values "good" or "bad", while a more sophisticated set might contain values "rated R for nudity", "rated R and recommended for minimum age 15", etc.

Guided Tours

All sorts of guided tours and independently threaded super-structures can be readily realized within the generic architecture outlined above.

For example, we have set up a number of tour sets which give different tours through a single document collection, namely the *WebLouvre museum*. There is one tour about the Baroque, one tour for the French painters, and a tour for impressionists. As an example, a page describing the work of Claude Lorrain is both on the Baroque tour and on the `FrenchPainter` tour. But when a user has selected a certain tour (by activating the corresponding annotation set), each page contains only the relevant navigation signs: If we look at Lorrain in the context of the Baroque tour, then the sign will lead us ahead on this tour; if we are on the `FrenchPainter` tour, it will point to the next tour stop on that tour. We can control access and/or charges for tours independently, and there can be any number of such tours--no-one will be confused by a multiplicity of signs on a given page, and the guidance information scales with the number of tours.

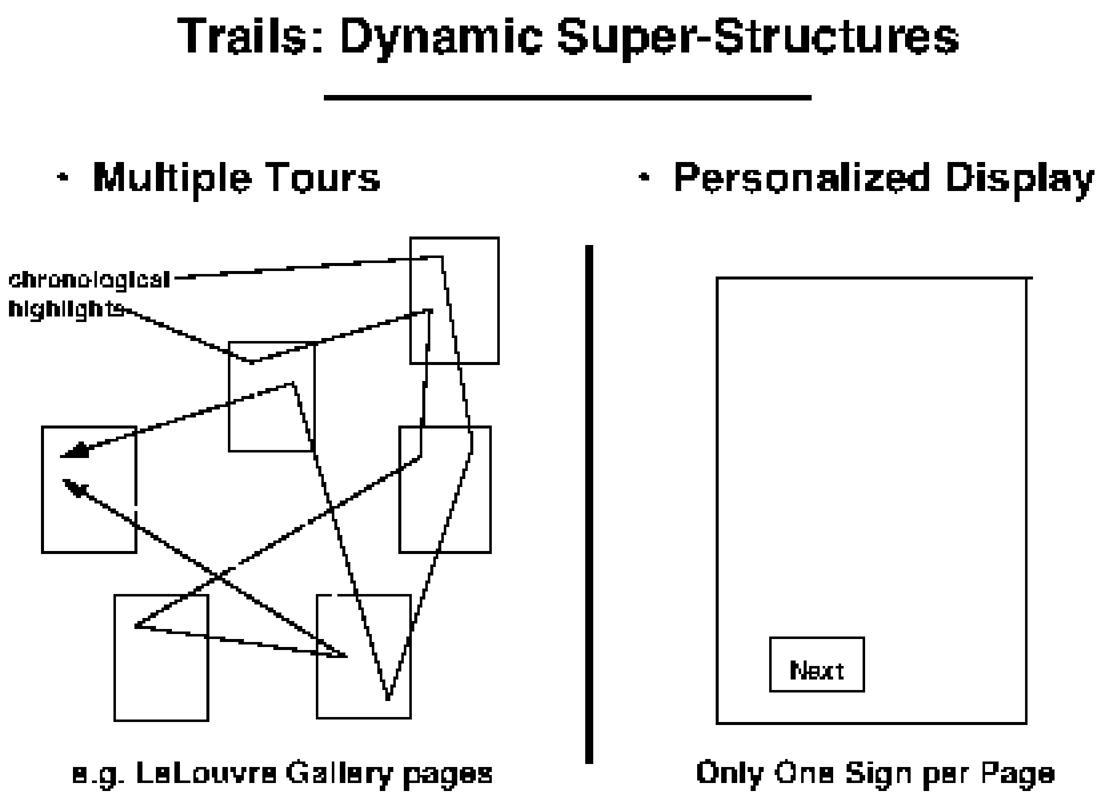


Figure 10

Conclusion

We have developed a generic mechanism for value-added third-party information, along with corresponding browser extensions for chosing perspectives, adding annotations, and adminstrating access control. This mechanism enables a number of usages which add value to the original information.

We have designed a scalable architecture, which distributes resources in a way that reflects social, economic, and legal boundary conditions. The associated protocol allows retrieval of access-controlled meta-information about documents uniformly and extensibly; it can be layered on top of existing protocols such as http. A prototype implementation, using a modified XMosaic browser and server scripts, has been tested and is being made freely available.

The architecture enables people to maket heir expertise and comments available to a selected audience in a structured way. We would like to see it contribute to a culture of widely distributed commenting and reviewing, much like the Web has led to a culture of individual publishing.

References

[BUSH]
Vannevar Bush (1945). Stanford Integrated Digital Library Project. Computer Science Department, Stanford University. (NSF/ARPA/NASA proposal). URL: http://www-diglib.stanford.edu/diglib/.

[COMMENTOR]
Martin Röscheisen, Christian Mogensen, and Terry Winograd (1995). ComMentor Documentation. URL: http://www-diglib.stanford.edu/COMMENTOR/.

[TR]
Martin Röscheisen, Christian Mogensen, and Terry Winograd (1994). A Platform for Third-Party Value-Added Information Providers: Architecture, Protocols, and Usage Examples. Available at URL http://www-diglib.stanford.edu/COMMENTOR/.

[TimBL]
NYTimes (1995). New Internet feature will make voluntary ratings possible. Proposed self-rating extension from Tim Berners-Lee.

[Borenstein]
Borenstein, N. (1995). KidCode. Draft. Available at URL ftp://ftp.fv.com/pub/nsb/draft-kidcode.ps

---

The Digital Library

P C D