WorldView: A System for Information Merging and Sharing

Pedro Furtado (pnf@dei.uc.pt)
Paula Gonçalves
Paulo Ferro
Henrique Madeira

Abstract

The WorldView system proposes a new way for merging and sharing information in workgroups for educational purposes. It offers universal web browser access and complete user interactivity based on interesting information structuring concepts and access permissions. It supports the common document and pictorial information but innovates by offering extended representation and interaction capabilities for users. This system shifts the focus from hypertext documents and hypermedia presentations to a workgroup system allowing interactive users to manage general information items. The user can insert or delete information items but he can also interactively edit, reference, relate and annotate any number of objects. These interactions are kept in a logical structural organization with individual user, workgroup and generic folders. Authenticated users can view, insert, modify or delete objects over the internet relying on an account and security administration. A database engine is used to store, index and search the data. We take advantage of the best ideas from database storage and searching, unix accounting and security, windows directory structure, hypermedia anchoring and linking and some structuring and object extensions of our own to offer a significant improvement in workgroups. The system can be deployed on its own or as a workspace system in any database repository.

1. Introduction

Information exchange and collaborative work are important domains of application for modern database systems [Silberschatz 91]. These applications became especially attractive due to the availability of the Web and the possibility of accessing a database from anywhere using a simple browser. We refer to systems with a large quantity of data that can be shared among people, especially semantically rich data such as pictorial data. Many interesting examples of such usage can be given. A virtual museum can keep images of art objects together with annotations and explanations about the objects. Digital libraries, satellite imaging , medical imaging databases are other examples.

A subset of those applications would benefit from group work and controlled user interaction over the information items so that authenticated users interactively build a information repository. WorldView is a research project in knowledge structuring and hypermedia tools for information merging and sharing. The system described in this paper resulted from the real needs of different professionals in our community to intuitively integrate and share a wealth of information in their group work and from the educational institutions needs to intuitively build and share educational material in their organizations. The common facilities from file systems with information retrieval capabilities, databases and hypermedia authoring tools are applied in an integrated way. The conjunction of database retrieval and hypermedia browsing allows the user to either search or browse and link the information. A prototype of the system will be used in a history department where student and research teams search for all available information on ancient monuments and paintings to reveal their history. They progressively insert their findings in workgroup spaces and decide what to share. A similar system could be deployed in a medical imagiology environment to insert and share image analysis and diagnostics.

WorldView offers a base structure to guide users organizing the information they insert and comprehensive user permissions and object privilege features to allow them to control the access in the web-based system. The basic features of document and image viewing and linking are extended to more general objects called information items. Information items can be inserted, deleted or edited depending on authentication. When integrated into the repository, pieces of documents or pictorial data can become interesting objects on their own. This is called information items within-component referencing. These components or information items can be related to other objects in the database possibly originated from other data sources. . The system allows users to reference, link or annotate any information item. The information is all kept in a directory-like structure. The system deals with the necessary security issues and establishes groups to enable users to choose the scope of the audience for their information. It is scalable to serve a community of any size and provides an extended database system which can be applied on top of a regular schema database with application in, among others, the educational, scientific, medical and art domains. The paper is organized as follows. Section 2 shows related work, Section 3 describes the system and Section 4 makes some commentaries about the application prototype. Section 5 concludes the paper.

2. Related Work

The system described here conjugates the philosophy of hypermedia collaborative work with complex semantically rich databases. Hypermedia systems organize and access multimedia information in a very intuitive way, linking all types of information. For a possible taxonomy of Open Hypermedia Systems please refer to [Osterbye96]. Hypermedia models such as the HAM model [Campbell 88] or the Dexter model [Halasz 94] have helped laying the foundations for these systems. For example, the HAM model divides those systems in three functional layers: a storage layer providing persistence, an application layer providing the functionality and a presentation layer for user interaction. Nodes, links and composites are important components of those systems which allow complete information linking. A database engine is used in many of those systems to efficiently implement the storage layer. Systems like HyperStorm [Bapat 96] or HyOctane [Buford 94] use an OODBMS to render hypermedia functionality. A general application framework for hypermedia documents on the VODAK ODBMS of HyperStorm can be seen in [Wasch 95]. WorldView extends typical hypermedia systems features by providing a base structure and access policy to allow users to freely create and integrate information and provides a web browser based access to users.

3. Overview of the System

The WorldView project builds applications incorporating many distinct parts that are assembled to work as a whole. It encompasses the following parts also illustrated in figure 1.

- Users/Groups
  ⇒ *the persons accessing the system. They form workgroups.*
- User Accounts & User and Object Privileges
  ⇒ *logged users can manage their objects and those from groups they belong to.*
- User Interface and Access Structural Organization
  - folder directories
  - user/group/generic workspaces
  - …

  ⇒ *this flexible system provides an important base structure to organize information.*
- Extending Metadata Objects
  ⇒ *referencing and Relationship objects, their specializations. Also Annotating objects.*
- Information Items & Schema
  ⇒ *documents, images, arbitrary pieces of both, objects from an existing database schema and database objects inserted by the user.*
- Database Storage & Searching
  ⇒ *a database engine implements the actual storage and indexing.*
- Support for External Files & Handling
  ⇒ *information can be stored externally or in the database. SGML documents can be further converted to corresponding document image objects and back to files.*

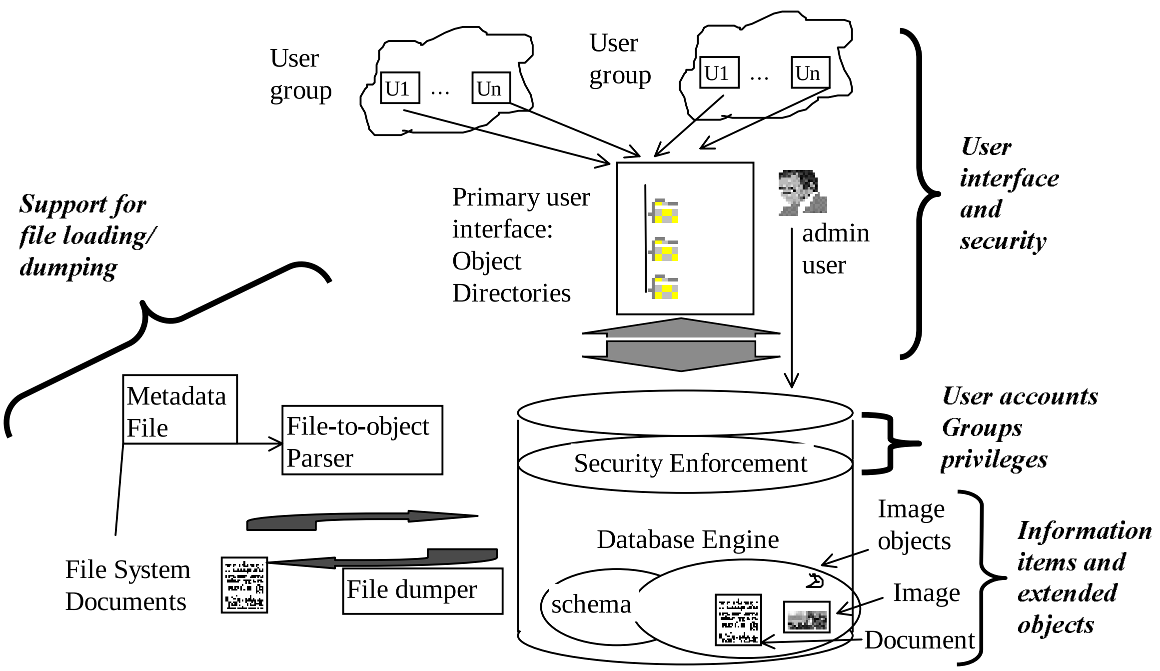


Figure 1- Overview of the WorldView System

The information items and the corresponding database storage and searching capabilities provide the basic functionality to store and retrieve documents and images and to reference external documents. SGML documents can be converted to document image objects by the external files handling part and back to files using a metadata format file. The schema represents possible particular application domain objects. For operation purposes, the information items, schema and extended objects are handled in the same way.

From now on we focus on the new significant parts which provide the facilities for user interaction.

3.1 Information Items

Information items are pieces of information like documents, images, document components (chapters, sections, etc) or video frames. They can be stored as database objects, their URL simply referenced or stored and referenced from the file system. The system offers the tools for referencing, relating and annotating those information items. It also provides the means to interactively define and reference within-components of information items, transforming those components into information items on their own. Image objects are within-components of images and their spatial coordinates or contours build referencing objects. Document anchors are another example of referencing objects. Sets and lists are also important objects. Sets are used to implement the WorldView folders which are the basic structural unit. Lists will be used in the future to create another type of folders called “sequence folders” which keep information items in a specific sequence. Besides those object types, users can also create database objects and relate those objects or objects from a preexistant schema to the information items.

3.2 Data Storage

The information items are stored in a database schema. Images and documents can be stored as BLOBs or simple references to external documents or images can also be stored. The Annotation is another base type. To achieve hypermedia and extended functionality, referencing and generalized relationship types have been added. The referencing object type is further specialized for referencing specific types. These specialized object referencing types include image region contours, image location pointer, text region and text anchor. The generalized relationship type relates any number of objects. One of its specializations is the familiar Hyperlink, but it also relates any number of objects. These relationships are useful when a user is viewing an object and wants to access the objects which are related to it, generalizing the hyperlinking mechanism to any objects.

3.3 Access Interface and User Security Enforcement

The access structure plays an important role in WorldView. The interface configuration is based on the following premises:

- ⇒ Information items are placed in folders.
- ⇒ Providing some structuring enforcement and at the same time allowing a reasonable degree of information items handling flexibility.
- ⇒ The system can be accessed by anyone through a web browser.
- ⇒ The access structure must be associated with a user acounting and security policy

Figure 2 illustrates the base structure. Registered users organize information in nested folders according to groupings, subjects or whatever is important. There are three base folders corresponding to possible useful repositories. The “all” folder is a generic workspace for everyone. The “group” folder contains any number of restricted access group folders for group workspaces. It is optionaly administered by a user with all privileges over the specific group. Finally, the “user” folder keeps individual optional workspaces for registered users. The folders marked with an “S” label show only one type of information item for direct retrieval of information of a specific type.

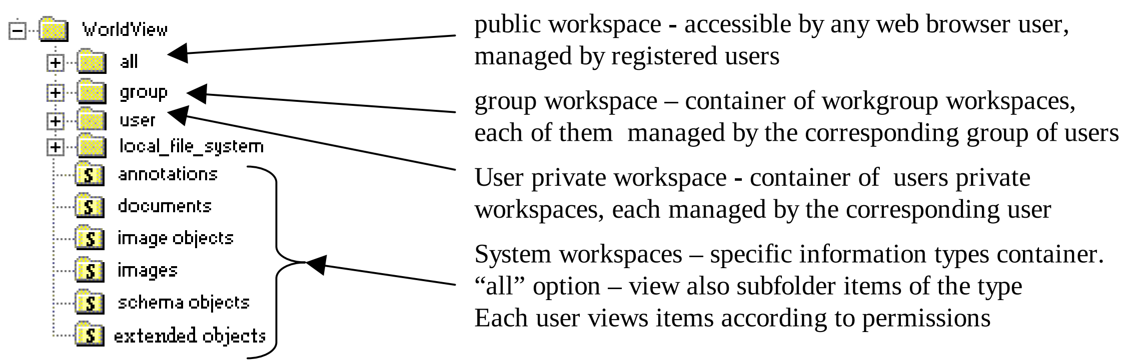
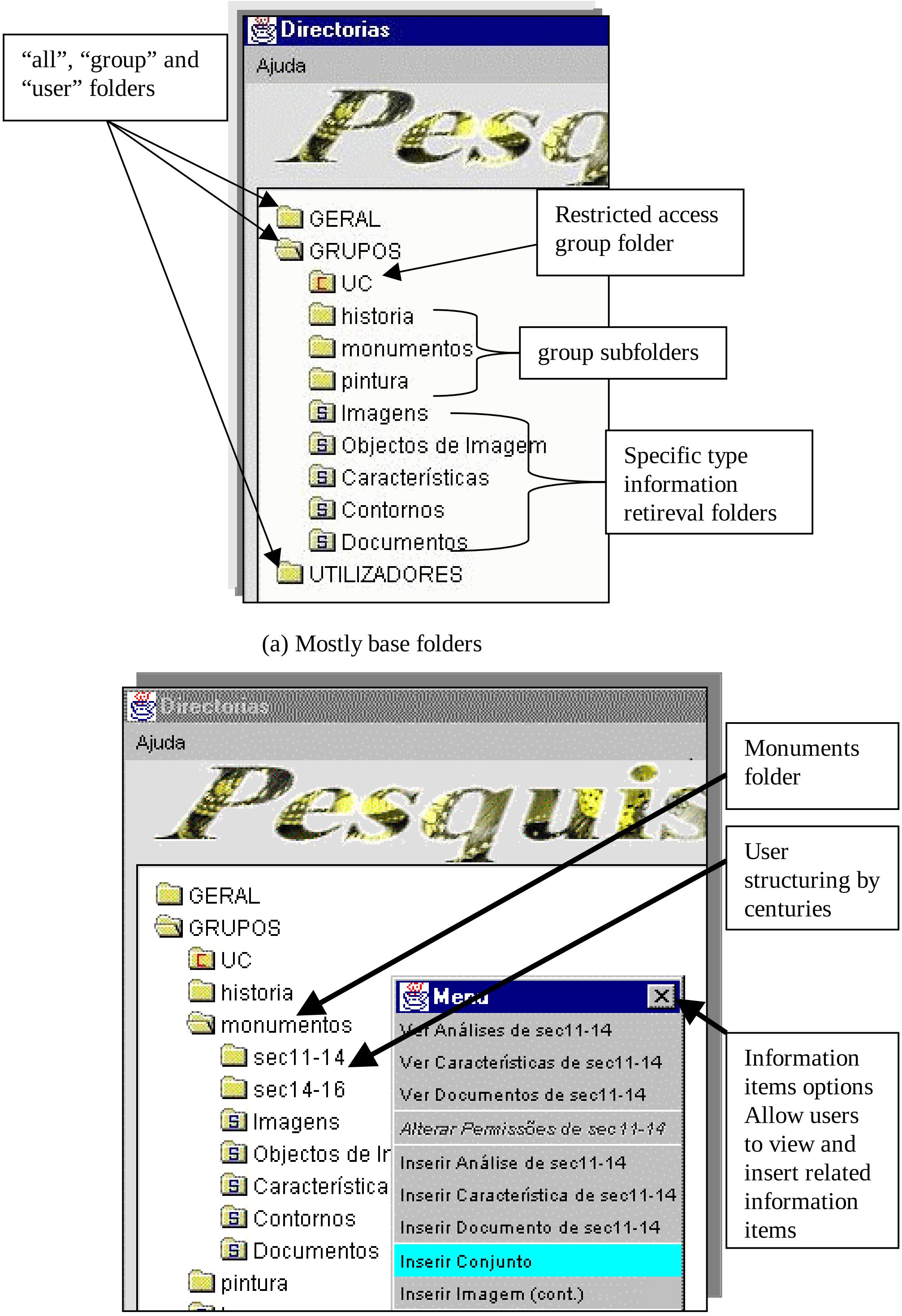


Figure 2- WorldView root folders

Specific groups can only be created by the administrator. Figure 3 shows some group folders access in the prototype implementation[1]. Each group can be public or restricted by default, but members can always access and extend them. User workspaces are private by default.

[1] The protoype is an adapted version for historical monument and painting studies. The examples are displayed in Portuguese but the important points are annotated with an English legend to clarify their meaning.



(a) Mostly base folders

(b) user nested folders

Figure 3 –Screen captures showing group folders

Users can assume several roles, depending on authentication. There are unlogged users, registered users, logged users, the administrator and specific group administrators. Users accessing the system are "unlogged users" until they submit a username and password, which they can do at anytime. A new account proposal is automatically mailed to the administrator for the first login. A user becomes a registered user if he owns an account. By default he will have a certain quota which is adjustable by the administrator. If the user has an account and enters a valid login he becomes a "logged user". Logged users can belong to many user groups (group administrators evaluate the applications). Every object has a owner (the creator), a group and a set of access permissions for user, group and other. Only the owner and the administrator (or group administrator) can change those permissions. Figure 4 illustrates the notation for object permissions. Access refers to viewing (which includes opening a directory), change refers to editing or modifying. Extending refers to inserting new objects, referencing, relating or annotating and the meaning of delete is immediate.

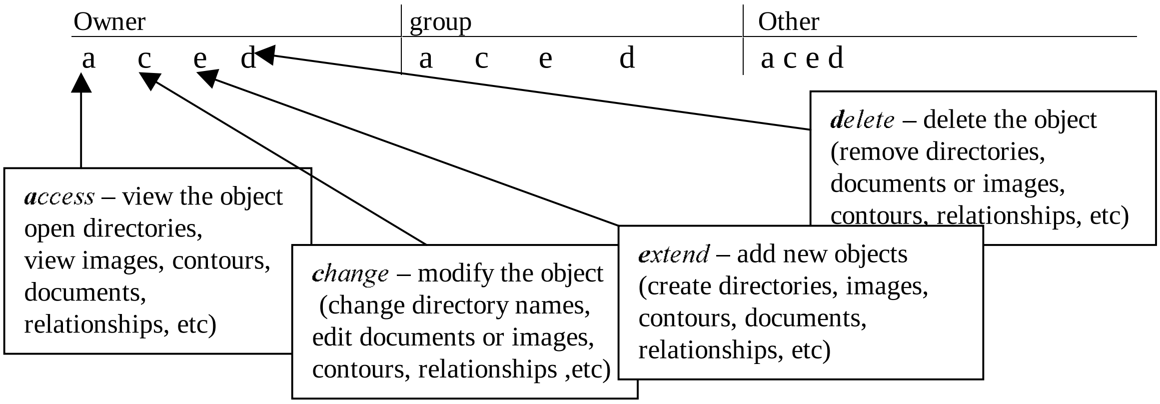


Figure 4 - Object permissions

Default object privileges are applied according to their workspace, as illustrated in figure 5. These defaults are changeable by the administrator and both the administrator and owner of an object can change its privileges. Unlogged users do not belong to any group, so they have "other" privileges. According to default privileges, they can access the information in the "all" workspace but cannot change anything. Public information can be put there. Logged users can extend the workspace by inserting information items and referencing, relating and annotating objects, but only the owner and the administrator can delete objects.

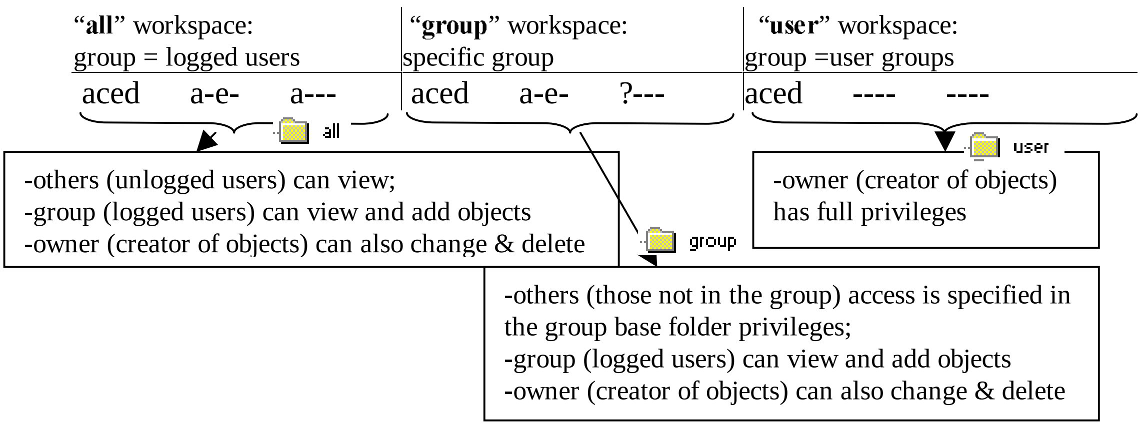
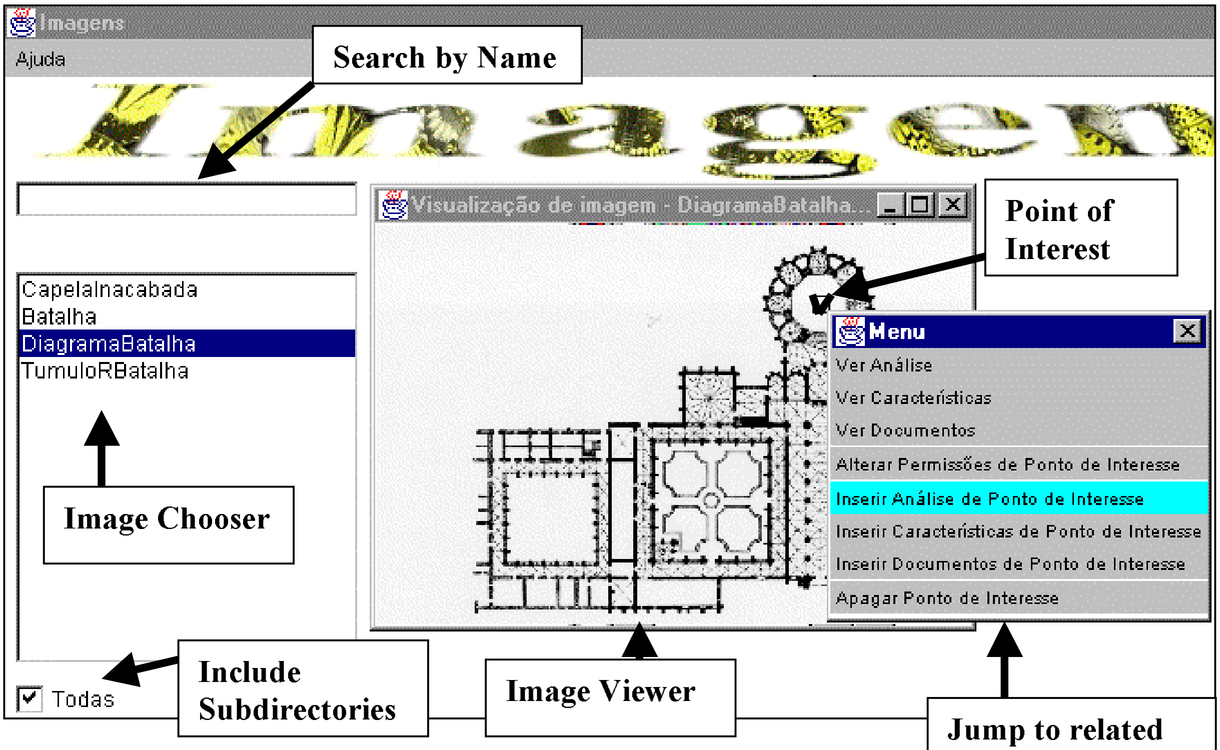
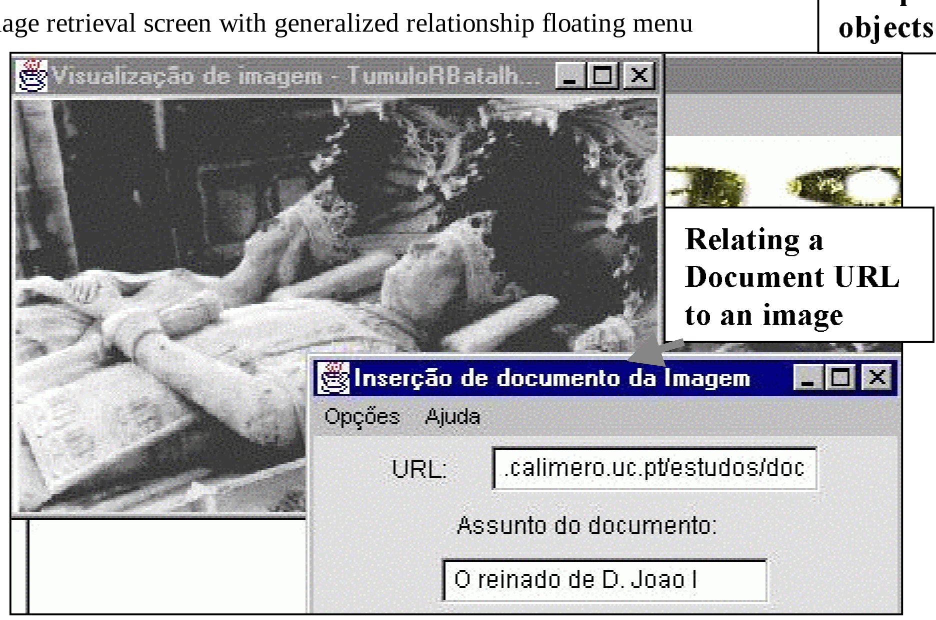


Figure 5- Default Objects permissions

Figure 6 shows some interesting features of the prototype application in what concerns the handling of information items. Figure 6(a) shows the image retrieval screen (there are similar screens for the other information item types like documents, annotations, image regions, etc. It allows the user to search for specific images, view one or more images simultaneously, annotate images, relate them to any other objects, create points and regions of interest and relate them to other items. The user can also view related information items.



(a) Image retrieval screen with generalized relationship floating menu



(b) Linking an image with a related document

Figure 6- Examples of retrieval, browsing and relating

4. WorldView Applications

The WorldView project is currently under progress. From it resulted a first workgroup application to share the material from studies of ancient monuments paintings and history in our country (this application was used to capture the screens in previous figures). The structuring offered by the system and the security implementation makes it the perfect tool to deploy in the school. A student has access to the historical documentation and research in progress using a simple browser but does not change or extend it unless he belongs to a workgroup. The researchers are grouped by their specific studies and therefore are able to view everything but can only extend and change information in groups to which they belong. This application uses only some features from the system and is centered around images.

The administrator installs the WorldView data schema by running an SQL script in any database engine of his choice. The JAVA application then connects to it by means of a JDBC string and all accesses to the database are made through standard SQL. This way WorldView applications are easily deployed. Users simply access it using a web browser. The current prototype is running on a Oracle database, but by changing the JDBC connect string it runs on top of any other database engine.

5. Conclusion

This paper presents the WorldView project. The system proposes a convenient way to share information with a universal interface accessible through a simple web browser. It presents a innovative user interface with information organizing structuring elements, user authentication, user and object privileges. It uses any database engine as its storage layer, which is accessed in standard SQL. The system provides the means to integrate any kind of information items, including within-components referencing and relating any number of items. A first application is shown which will be used to support research studies of ancient monuments, paintings and history in our country. The research and implementation work is still under way to build a system supporting all the intended features. We would like to further research the inclusion of information items sequence grouping capabilities so that the user can be directed to navigate a sequence of items.

References

[Bapat 96] A. Bapat, J. Wasch, K. Aberer, J. Haake "HyperStorm: An Extendable Object-Oriented Hypermedia Engine" in proceedings of the 7th ACM Conference on Hypertext, March 16-20, 1996.

[Buford 94] Buford J., Rutledge L., Rutledge J. and Keskin C., " HyOctane: A Hytime Engine for an MMIS", Multimedia Systems (1)4, February 1994.

[Campbell 88] B. Campbell, J. M. Goodman,"HAM: A general purpose Hypertext Abstract Machine", Communications of the ACM, Vol. 31, n 7, July 1988, 856-861.

[Halasz 94] F. Halasz, M. Schwartz, “ The Dexter Hypertext Reference Model” in Communications of the ACM, Feb. 94, Vol 37, n 2.

[Osterbye 96] Karper Osterbye, U. K. Wiil, “The Flag Taxonomy of Open Hypermedia Systems” in 7th ACM Conference on Hypertext Online Proceedings, Washington, USA, March 16-20, 1996.

[Silberschatz 91] A. Silberschatz, M. Stonebaker, J. Ullman eds. “Dtabase Systems: Achievements and opportunities” in ACM Communications 34,10, October 1991.

[Wäsch 95] Jürgen Wäsch, K. Aberer: Flexible Design and Efficient Implementation of a Hypermedia Document Database System by Tailoring Semantic Relationships: GMD Technical Report No. 908.