ACCESSING THE COLLECTIONS OF AN ENGLISH COUNTY THROUGH TIME, SPACE AND THEME

C. DOBBYN, M. ZAJICEK AND R. PHILLIMORE, SCHOOL OF COMPUTING AND MATHEMATICAL SCIENCES, OXFORD BROOKES UNIVERSITY, UK

ABSTRACT

This paper describes ideas developed in the specification and outline design of a multimedia distributed system (ORION), completed by the authors under contract to Oxfordshire County Council, Department of Leisure and Arts. The system specified will make multimedia representations of items in collections in the care of the County Council available and easily accessible to the community. We describe our concept (VECTIS) of a system that enables users to explore different themes across the history and geography of the County of Oxfordshire. The interface design is innovative in that users can explore the topography of the County by traversing 2-d and 3-d maps—zooming in on areas of interest—and focus on, or travel across, specified time spans using a timeline. At any stage, the system is able to retrieve details of those items across all of the Department of Leisure and Arts' collections that are relevant to the time, place and theme the user is currently focusing on, in the form of images, video, audio, text, etc., through integrated indexation of a set of widely different collections (museums, archives, natural history collections, etc.) Users may then examine multimedia representations of items that are of interest to them through appropriate viewers.

KEYWORDS

agents, distributed database, space, theme, time

BACKGROUND

The County of Oxfordshire in the United Kingdom holds magnificent collections of books, documents, relics and artifacts relating to its rich topography and long history; but it is reasonable to say that the majority of Oxfordshire residents know little of these collections. Those who are aware of them and wish to examine any item—and know where to look—have to travel to the different centers where individual collections are stored. There they must often rely upon the expert knowledge of specialist staff to help them gain access to the specific items they wish to see, to interpret these items for them, or to direct them, when necessary, to other collections in different locations.

Given the huge interest in, and growth of, multimedia information handling over the Internet, and with the increased bandwidth of digitized traffic over cable and satellite connections, it was clear to the Oxfordshire County Council *D*epartment of *L*eisure and *A*rts (DLA), the custodians of Oxfordshire's historical and scientific collections, that the future must lie in some form of distributed multimedia system. In late 1996, the School of *C*omputing and *M*athematical *S*ciences (CMS) at Oxford Brookes University were approached by the DLA with a view to developing, in partnership, a specification and outline design for such an information system, to be submitted to the Heritage Lottery Fund as a proposal for funding for its development. The main aim of the system was to further the DLA's mission of *Access*, *Integration* and *Conservation* by making more widely available to the community the treasures which are held by the council on their behalf: digital records of photos, artifacts, printed materials and archives items, etc. would be made, and these records stored in a multimedia database with distributed access. Above all, the system was, by making available representations of the County's artifacts and its history to its people and to visitors, to foster a *sense of place*.

More particularly, the constituents of the system were to be:

- a search and query system of a novel and intuitively usable nature;
- systems to search for items across indexes of all Oxfordshire collections;
- appropriate hardware and software for retrieval and display of multimedia representations of items in collections;
- generic software systems encapsulating the expertise involved in certain common areas of en-

© ARCHIVES & MUSEUM INFORMATICS, 1997

quiry (e.g. genealogy), to be used as automatic information gathering agents, or to teach appropriate information-gathering techniques;

- software to encapsulate material and data which will enable local studies and genealogical research in Oxfordshire, with a view to sale on CD-ROM;
- facilities to set up virtual 'guided tours' through the collections, based on selected themes (e.g. transport, a specific town, etc.);
- a means of configuring the indexes of the system to introduce new themes, categories, etc., without re-programming.

It is worth indicating here just how big a system was being proposed. There are nearly 3 million records in the Oxfordshire collections. Of these, we estimated that about 1 million were suitable for digital capture—artifacts, photographs, maps, documents and ephemera. The storage space needed for such a collection of records was estimated as about 750 GBytes; the labor, alone, in capturing these representations would be 600 person-months, at a cost of over £1 million. Clearly, powerful networking capabilities and a sophisticated human-computer interface would be required for such a huge store.

The specification for ORION (the *O*xfordshire *R*egion *I*ntegrated *O*n-line *N*etwork) was completed on time and submitted to the DLA in January 1997. The main body of this paper describes the system we have proposed: we lay most emphasis on the details of the principal user interface, a search and query engine based on the concept of *V*irtual *E*xploration of the *C*ounty in *T*ime and *S*pace (VECTIS). However, later sections deal in outline with the design of the search and integration software and with the global indexes and hardware that enable the system to function.

VECTIS—Exploring Oxfordshire in Time, Space and Theme

Introduction

Oxfordshire is a large English county whose landscape bears countless imprints of the humans who have inhabited it since prehistory, and who have left behind them myriads of texts, images and artifacts as a testimony to their passage. The earliest question the CMS development team had to answer was this: In what ways might interested parties want to investigate this storehouse of information?

There can, of course, be no definitive answer to such a question: people approach the County's archives and collections with purposes as varied as people themselves. Doubtless, certain users have a very clear idea of what they are looking for: they might want a reference to a particular source, whose author and title they know; or they might want to see a specific artifact; or they might have a particular site in mind to visit. It is easy to devise schemes of query and access for such researchers, and we propose one such graphical scheme in the specification. But the evidence of the professionals who are responsible for museums and study centers suggested that many—a majority probably—of public enquiries are of a general nature: clients wish to know more about a particular area, perhaps; or seek information about a building, a company or an institution over time. Certain sorts of investigation are so common as to be generic; two, at least, come up time and time again: research into family history and project work on some specific theme, such as transport, or elections. Some visitors have only a vague idea of what they want to see; some just come to browse. Unfortunately, however, the team were given neither the time nor the financial resources to set up the prototypes and empirical investigations necessary to provide a conclusive answer to the question of what the more general user is seeking in approaching Oxfordshire's collections. Was it possible to impose a structure on these many and diverse types of interest?

We recognized that there were two dangers in devising a system through which clients can explore and record the County's store of information. The first danger—a common one—was of specifying a system with structures which simply mirrored the schemata and retrieval processes of the underlying computer storage, and which would thus be difficult for ordinary members of the public to use. The second hazard was of devising exploration processes based on a natural and human view of the County's information, matching a user's own mental models [Carroll 1987; Monk 1984; Day & Kovacs 1996]—which would hence be easy to use—but could not cater for certain common types of query. We decided that these possibilities could be guarded against in three ways:

1. Searches and enquiries should be *graphically based*; this is in keeping with the design of more or less every modern software system [Brown 1988; Bertin 1981];

2. The interface to the underlying data structures—the catalogues, databases of images, computerized archive materials, and their indexes—should be based on natural concepts independent of these structures;

3. The system was to be to some degree *configurable* by DLA professionals, based on their experiences of how the system is actually used, without any re-programming and re-compilation. The system was also, to a limited extent, able to *learn* from the explorations it is asked to undertake, that is, it will be *self-configuring.*

In the next section, we consider the three fundamental dimensions on which we propose that users may search DLA information systems: *Space*, *Time* and *Theme*; and in the section after that we discuss how users may maneuver in these dimensions.

THREE DIMENSIONS OF SEARCH

SPACE

Consider a 2-dimensional map of Oxfordshire. Consider, further, an enquiry at the Center for Oxfordshire Studies, say about some very specific point in the County—an example might be "what is Wayland's Smithy?" We can represent the area of space examined as a single point on the map, as in Fig 1. Now suppose the hypothetical client has a more general idea in mind: she wants information—though at this point may be uncertain as to exactly what information she wants—about archaeological sites around the Uffington area. We can represent the client's spatial area of interest as a square (perhaps), of an uncertain size (who knows what "around the area of" means precisely?), on our map, as in Fig 2.

But a client may not be in a position to specify a query in some precise geographical area. Their area of interest may be some general theme covering the whole county—the Roman roads of Oxfordshire, for example, in which case the entire area of Oxfordshire is the spatial domain of the research. More interestingly, a possible investigation might be on a topic that implies a number of fairly specific areas of the County whose relationship to one another is that theme. Our client might, for instance, be interested in a certain company and its effect on its local community, that company having had several branches around the County. In such a case, the spatial aspect of the query would resemble Fig 3. The client might not know, of course, of the existence of all these branches; and it would be a responsibility of the system to fill in these details for the client, either by default, or as part of a spatial broadening process. Indeed, the concept of *broadening* in space, and in the two other dimensions of exploration discussed below, is an important one in the design of the user interface to this system; we believe is an important contribution to the DLA's prime goal—that of a sense of place. The concept of spatial and other types of broadening are discussed in subsequent sections.

TIME

If we just provide facilities for clients to investigate Oxfordshire by specification of spatial areas, we are

Figure 1

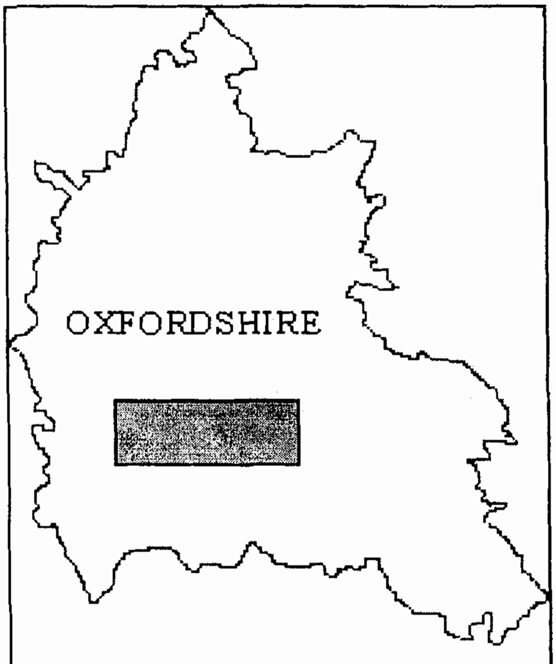

Figure 2

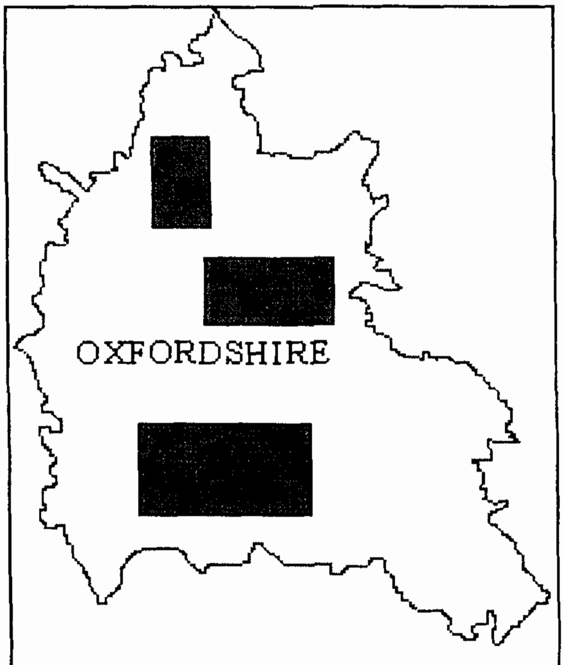

Figure 3

leaving out what is perhaps the most important dimension along which they might wish to investigate—that of *time*. Oxfordshire's collections are principally records of the history of the County—from pre-history to the present day. Returning to our 2-d map, we can now imagine it extended in a third dimension, representing time. Researchers enquiring, say, about the history of the church in Bletchingdon, are asking for a description of records within a cylinder cut through the 3-d space-time map, as in Fig 4.

Clients with more complex spatial queries, as described in the previous section, will find themselves with a set of such space-time cylinders. It should even be possible to specify a set of areas over a set of time periods. A user who is interested in records concerning churches in certain specified areas of South and West Oxfordshire during the Reformation, and also from 1900 to the present day, for example, would be in the situation depicted in Fig 5.

The concepts of broadening narrowing and traveling are, of course, relevant in the time dimension as well; and these are developed below.

THEME

But users will certainly not simply wish to enquire generally about certain areas of Oxfordshire: professional experience shows that a majority of users will have some reasonably specific theme on which they are looking for information. A client might, for example, be looking for information on *Oxfordshire transport*, or *Oxfordshire churches*. A theme may, of course, also be constrained by selections on the other two dimensions, space and time, such as in queries about churches *in South Oxfordshire, between 1800 and 1900*, or transport *in Oxford from 1900 to the present day*. The system will provide facilities for a user to develop and select such themes. The idea of broadening, narrowing and traveling is relevant to this third dimension of search, as well; and these concepts are elucidated below.

Themes can be understood as hierarchically organized, that is to say a broader theme such as, say, TRANSPORT, may have sub-themes—for example, PATHS and VEHICLES. These, in turn, may have sub-themes themselves, such as FOOTPATHS and ROADS. This schema, as with all such schemas, is a simplification of the conceptual structures we actually carry around in our heads; but such simplifi-

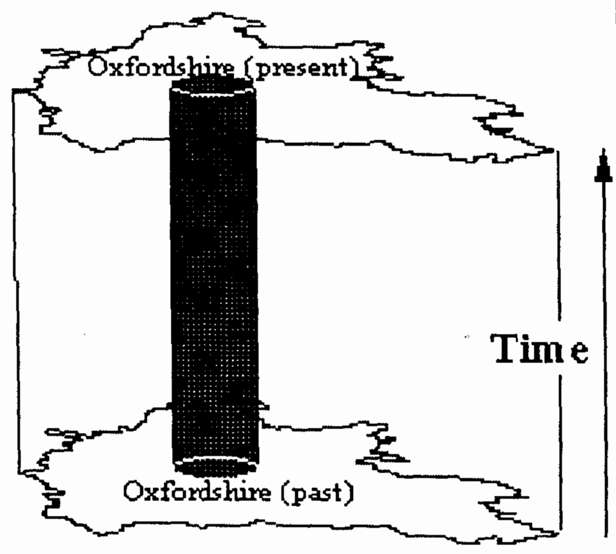

Figure 4

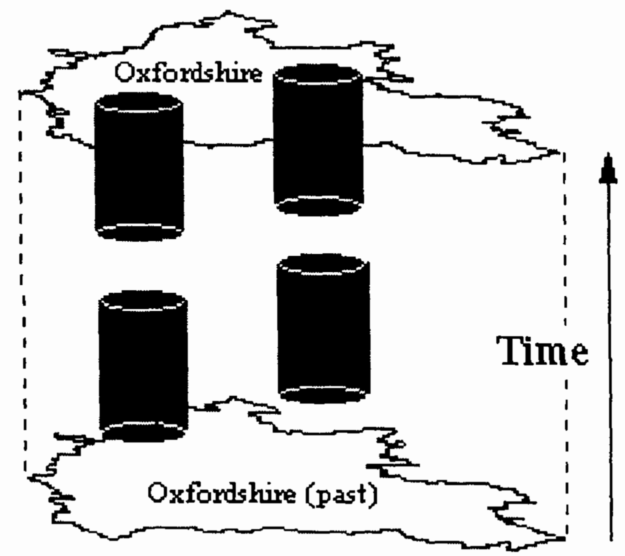

Figure 5

cations are always necessary wherever computers are involved. The derivation of a full thematic hierarchy will require a substantial effort of analysis.

Information Space and Focus

We can now picture all the DLA's records and other collected items as lying within a three dimensional *Information Space*, the dimensions of this space being space, time and theme. We can go further and portray this space as a 3-d cube, the interior of which represents all possible combinations of space, time and theme. Items in DLA collections thus occupy discrete points in this space. Of course, this is another simplification—items are persistent in time and many items may be classified under several themes; so a single item may occupy more than one point in the space. However, accepting this, we can now see that a client's area of interest will occupy some set of subspaces of the whole Information Space, inside the cube, as depicted in the Figure. We term this set of subspaces, the *Focus*. See Fig 6.

Broadening, Narrowing and Traveling

In space

Broadening in space implies that users should be able to increase the size of their area of interest, or—as in the example in above—incorporate new areas, which may possibly be spatially disjoint, but are related to a previously-specified area through some topic. Narrowing, obviously, means the converse: clients may decrease the size of their spatial area(s) of search by zooming in on some part of a previously specified region, or by excluding disjoint areas that had been selected previously.

Traveling along the spatial dimension means simply that users may select a starting point inside one of their specified areas and then move a detailed spotlight of interest in any spatial dimension, as illustrated in Fig 7. Users should be able to travel across both a 3-d and/or a 2-d representation of the area. Both representations would convey a reasonably detailed sense of the ground across which the spotlight is moving, and be marked with images of sites, landmarks or other places related to the current theme.

In time

Similar principles apply to the time dimension: clients might wish to increase the length of time over which they are searching to take in more information; they might also wish to add new, but disjoint, time periods. Conversely they may find that the time

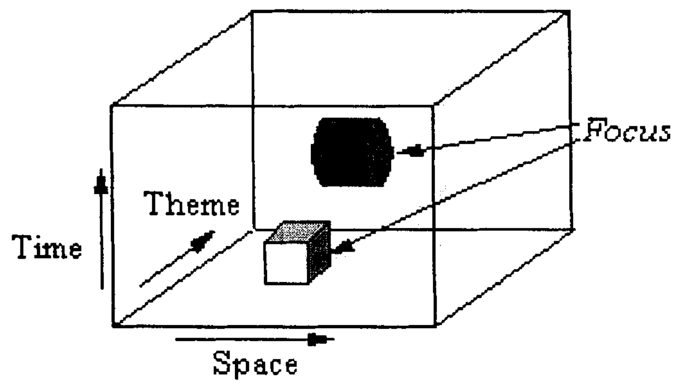

Figure 6

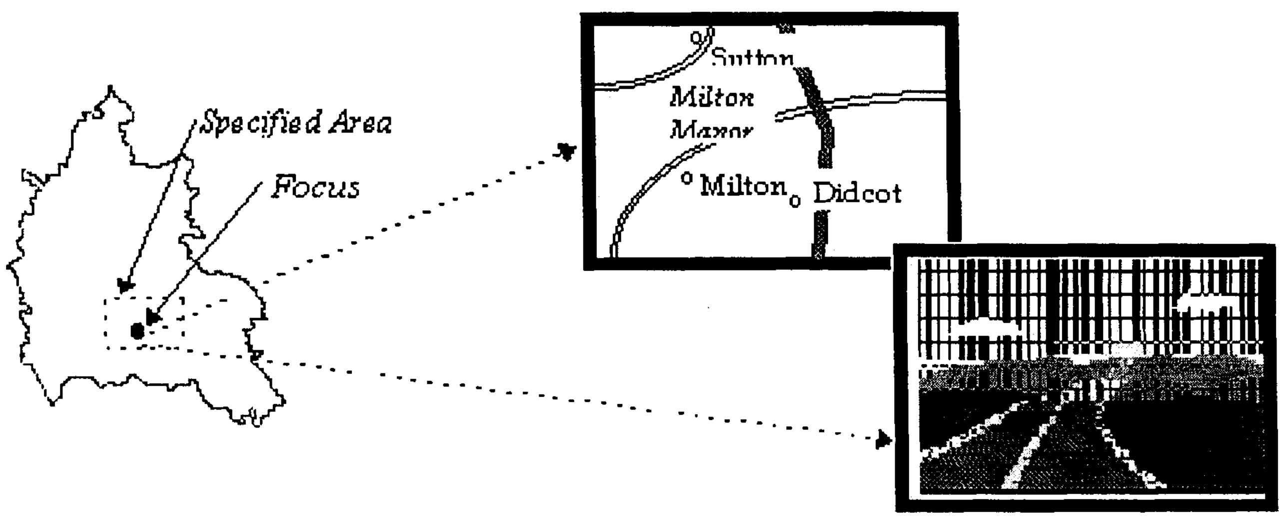

Figure 7

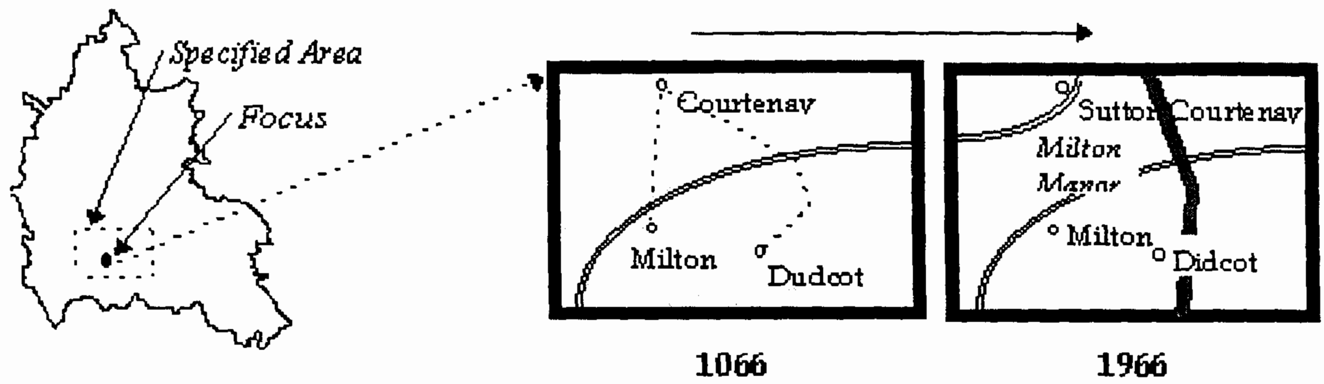

Figure 8

period they have specified simply contains too many records; and thus they will have to narrow the span of time over which they are seeking information. The system must contain facilities for such extensions and diminutions of search along the time dimension.

The concept of traveling in time should also be applicable in the system. Users may elect to move along the time dimension (in increments to be chosen by them). The system can then reveal 2-d or 3-d views of the specified area at each time interval, as in Fig 8, again illustrating features relevant to the current theme(s).

Naturally, as there are limits to the amount of data that can be stored, certain time jumps may yield little change. However, as in the illustration large leaps, or smaller jumps in eras of intense change, will show interesting alterations in patterns of settlement, ownership, technology and demography.

In theme

The concept of movement across the theme dimension is rather more elusive than that of movement in either space or time; however, there are clear-cut ways in which users can maneuver in this dimension by Broadening, Narrowing and Traveling. One way of narrowing a theme takes the form of *specialization*: selecting a sub-theme of the current theme; for example, a user might specialize an enquiry into the theme PLACES, by choosing the ARCHAEOLOGICAL SITES, a specialization of the previous theme. A broadening can then be seen as the converse: moving from a more specialized theme to a less specialized one; from COMPANIES, for example, to INSTITUTIONS. However, it is also possible to Broaden by incorporating new themes into a sphere of interest: a user currently focusing on the theme COMPANIES might wish to bring in some related theme, say PEOPLE, to be investigated concurrently

(see Fig 9 for a slightly different example). With this type of Broadening, the converse Narrowing will be the exclusion of a theme from the Focus.

We believe it is possible to travel in the Theme dimension, also, through a mechanism we call *Semantic Transition*. Users concentrating on a certain theme may move from that to a conceptually or semantically related theme. For example, if the current target theme is MONUMENTS then the system might allow the user to shift focus to the conceptually related themes of MEGALITHS or STATUES. We clearly recognize that no computerized

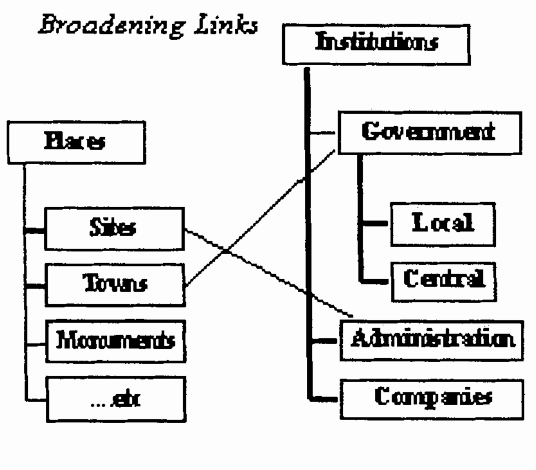

Figure 9

schema can possibly represent all imaginable conceptual links; so these links will be configurable by DLA users. The system can be set up to gather information about the theme transitions that users make, to be used as an empirical basis for the introduction of new conceptual links and the deletion of existing ones.

SELECTION

As users broaden and narrow their Focus in space, time and theme, the system indicates the number of records and other items that are contained in that area of the Information Space (this may be an estimate). The system can then be ordered to present a summary of the items available in that area of the space—usually, pursuing our graphical theme, in a choice of textual or iconic form; and clients can select some set of these for display at their terminal, if they exist in digital form, or for further details of where the items may be accessed if they do not.

The concepts outlined above—Space, Time, Theme, Focus, Broadening, Narrowing and Traveling—allow a form of interaction that we believe is both novel and natural, and which we have called VECTIS. We believe they form the basis of a system that is more than just an integration of databases, but one in which a sense of place can truly be created by allowing users to become travelers in time and space within the borders of Oxfordshire.

THE USER INTERFACE

This section describes briefly how the VECTIS concept actually manifested in the specified user interface to the ORION system; how ORION will look and feel to its public users, who come to it to retrieve information about Oxfordshire's cultural heritage; to explore its topography and history; and to browse through representations of the records and artifacts held by the DLA. We present, in mainly pictorial form, a description of some of the screens and images that might be displayed in a simple interaction and the facilities for manipulation and exploration of information that the interface affords its users. Essentially, the public interface has three components: the *Explorer*, the sub-system that incorporates the VECTIS concept, and through which users may traverse Oxfordshire's landscape and treasures past and present; the *Finder*, by means of which users who have some very particular need in mind may retrieve specific pieces of information; and the *Browser*, which allows users to leaf through information by conventional hypertext means. A number of *Generic Systems* are also provided which operate at a higher level: they carry out automatically certain common sorts of exploration, for instance a genealogical search for records of a certain family; these are not discussed in this paper.

THE DESIGN PHILOSOPHY

It is more or less a given that any modern software system requiring substantial interaction with human users will have a windowing interface [Schneiderman 1992]; and it is becoming more and more the case that this means Microsoft Windows. However, our aim was that ORION should reach as wide a public as possible; graphical interaction was to be an essential part of our design, and we accepted that any modern interface will incorporate some form of windowing. However, while preserving some of the spirit of MS-Windows interfaces [Microsoft 1992], our conception was based on the following key ideas:

- *Buttons*. Users would issue instructions to the system mainly by pressing buttons on the screen.

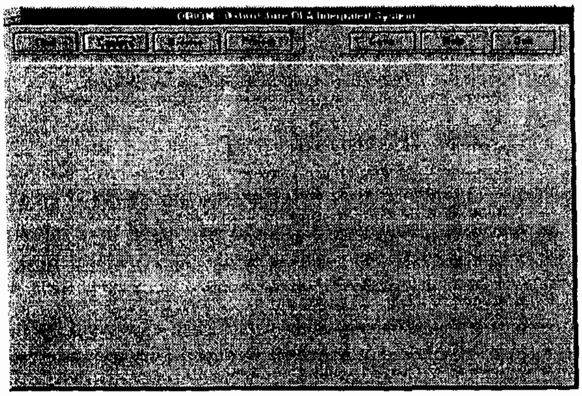

Figure 10

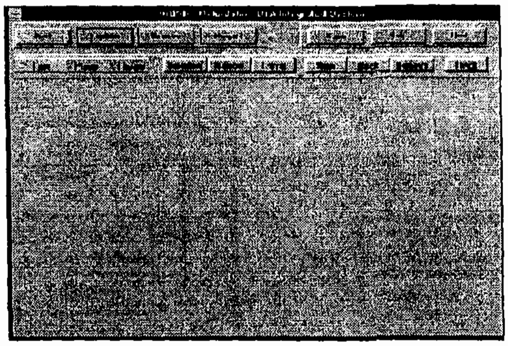

Figure 11

These buttons may actually be "pressed" with the mouse, a touch-screen, a joystick, or some combination of these;

- *On-screen Manipulation*. Other user wishes and preferences can be expressed by manipulation of objects on screen: users will move sliders, pull icons into areas, slide areas under viewers, etc.
- *Maps*. The spatial dimension of Oxfordshire would be represented by 2 and 3-dimensional maps, areas of which can be chosen, which can be moved over, zoomed in on, etc.
- *Time Lines*. The temporal dimension of the county will be represented by a set of time lines, familiar to us all. These lines can also be moved over, zoomed in on, etc.
- *Icons*. The Thematic elements of search are represented in terms of icons that can be dragged, dropped and otherwise manipulated by the user

We can now illustrate some of the ways in which users might interact with the ORION system. What we present is not the full specification of the user interface; it is a set of screens intended to illustrate the VECTIS concept and to give readers an insight into a simple interaction. Many details of precisely how a user might use the system—what buttons will be provided, how areas of screen can be selected, etc.—have been omitted for reasons of space.

The Main Screen

All elements of the ORION system are accessed from a main screen, elements of which remain visible throughout any user interaction. This screen is illustrated in Fig 10. The purposes of the three buttons on the right SAVE, HELP and QUIT, are more or less self-explanatory. SAVE commits the current state of the user's interaction to disc, to be returned to later—the prompting for a filename; HELP brings up a help screen and QUIT quits the session, prompting for a save, if required.

The four buttons on the left access the main elements of ORION, the Explorer (which incorporates the VECTIS concepts discussed above); the Finder, which allows the user to search for an item of which she already knows the name, or some other key detail; the Browser, which allows standard Internet-type access to HTML pages set up by the DLA; and the Generic Agents, which provide automated help with certain standard tasks, such as researching family history, local history, etc.

The Explorer

Assuming the user has pressed the "Explore" button, the Explorer sub-system is loaded and the Explorer tool bar becomes available; this consists of three check boxes, as illustrated in Fig11, which access the tools by which the three exploration dimensions can be manipulated, as described in the discussion above. The manipulation is carried out by means of a series of software agents, some of

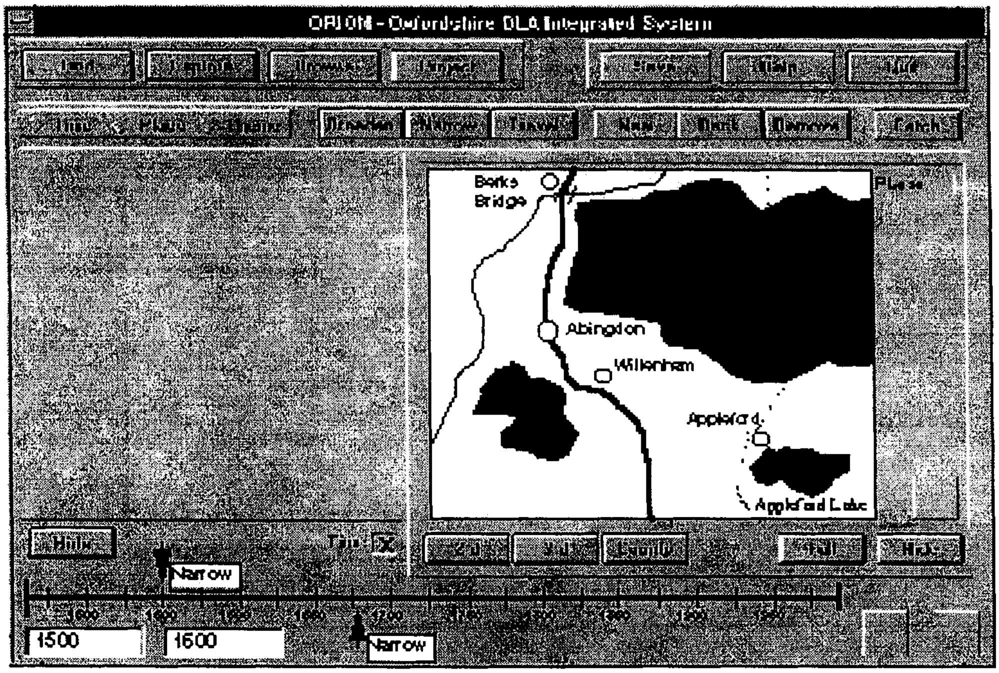

Figure 12

which have a visual manifestation at the user interface. These agents are briefly discussed in a later section; in this discussion, we are only concerned with three: the Space Agent, the Time Agent, the Theme Agent.

Manipulation in the spatial dimension is carried out by the Space Agent, whose task is to interact with users through a series of maps (the user may choose between 2-d and 3-d maps and alternate between these at will) and to interpret their inputs to the underlying system. The Time Agent has the visual manifestation of a series of timelines, on which users may broaden, narrow, etc. their focus of interest. The Theme Agent displays iconic representations of the hierarchy of themes and allows users to select themes, narrow the theme selected by specialization, move to related themes, etc. Fig 12 depicts the Explorer in a state where the Time and Space agents are active. Fig 13 shows a state where all three agents are active. Readers should note the interaction between the visual representations of the agents: if the user narrows her temporal focus to, say, the years 1500 - 1600, the maps displayed by the Space Agent will reflect the topography of the selected area at the end of that period; if the Theme is narrowed from, say, buildings to churches, the Space Agent will only display representations of the churches in the area of Focus (in the period specified) on its map. User interaction, in the case of all three agents, is carried out directly on the screen: in the Space Agent, areas of maps can be selected by dragging the mouse (or a joystick); in the Time Agent sliders are pulled along the timeline (and new timelines can be generated for temporally disjoint periods or for sub-periods); in the Theme Agent, icons can by dragged and dropped to new locations. All three agents can be manipulated in this way, interacting with each other, in the process of focusing interest on some area, time and aspect of the County's collections; this has been termed arriving at a *Focus* in earlier discussions. Once having found this Focus, users can now examine the items available in the DLA's collections that are subsumed in it. If the user presses the button marked "Fetch", the system will report as in Fig 14, the number of items that can be found and give users choices as to which, if any, to report further on. If the user opts to retrieve a limited number of records, there will be range of options as to decide on which to retrieve.

Those records that the user chooses for detailed examination in this way are bound together by type in a set of workbooks, which are presented through

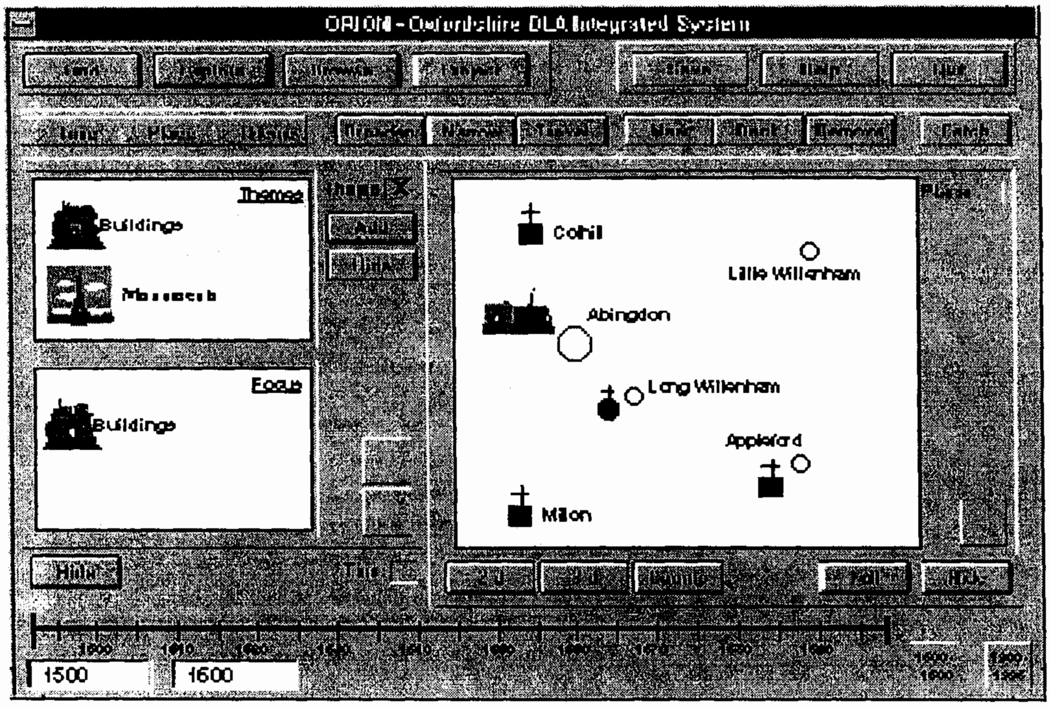

Figure 13

a set of viewers, appropriate to each record type. Fig 15 illustrates an 3-d Image Viewer displaying a workbook of 16 images (the user is currently examining the fifth). The user is able to flip through the book of images using the "Next" and "Previous" buttons, and carry out other operations on the image under display. The Explorer provides viewers for 2-d and 3-d maps, text, movies and other multimedia types. The reader will have noticed that the viewer illustrated displays a button marked "Interpret". If the user presses this button she will be supplied with further expert, interpretative information from the DLA's records regarding this record. This is illustrated in Fig 16: the upper pane of the presentation contains details of the further information that can be retrieved; the user drags the textual representations of these interpretative items into the lower pane; and this further information will be displayed in the pane, as depicted.

Certain DLA records (e.g. library books) will have no on-line, retrievable representation. However, ORION is able to supply sufficient information for these records to be retrieved physically by the user by going to the appropriate place.

THE FINDER

The Explorer, discussed above, is for casual users of the ORION system: tourists, students, members of the general public. However, as our earlier discussion indicated, many users may approach the ORION system with a very precise idea of what records they want to examine. These users are catered for by the second element of the ORION interface, the Finder. Fig 17 illustrates the main screen of the Finder. Users are invited to look for items within individual collections (e.g. Center for Oxfordshire Studies), within whole departments (e.g. Museums, Libraries) or across all the DLA collections, by checking the appropriate box and entering details of the item they are searching for; the user is not required to fill in all of these boxes: but a minimum of one is required. Fig 18 depicts the result of locating a known record. Details are provided of the title of the record: its author (if any, or any other person associated with it); the type of record (e.g. a book, an archaeological find, an archive entry, etc.); the reference number under which it may be accessed; and the location(s) at which the record may be found. The user is given the choice of retrieving an on-line representation of these items; but obviously this will only be ap-

© ARCHIVES & MUSEUM INFORMATICS, 1997

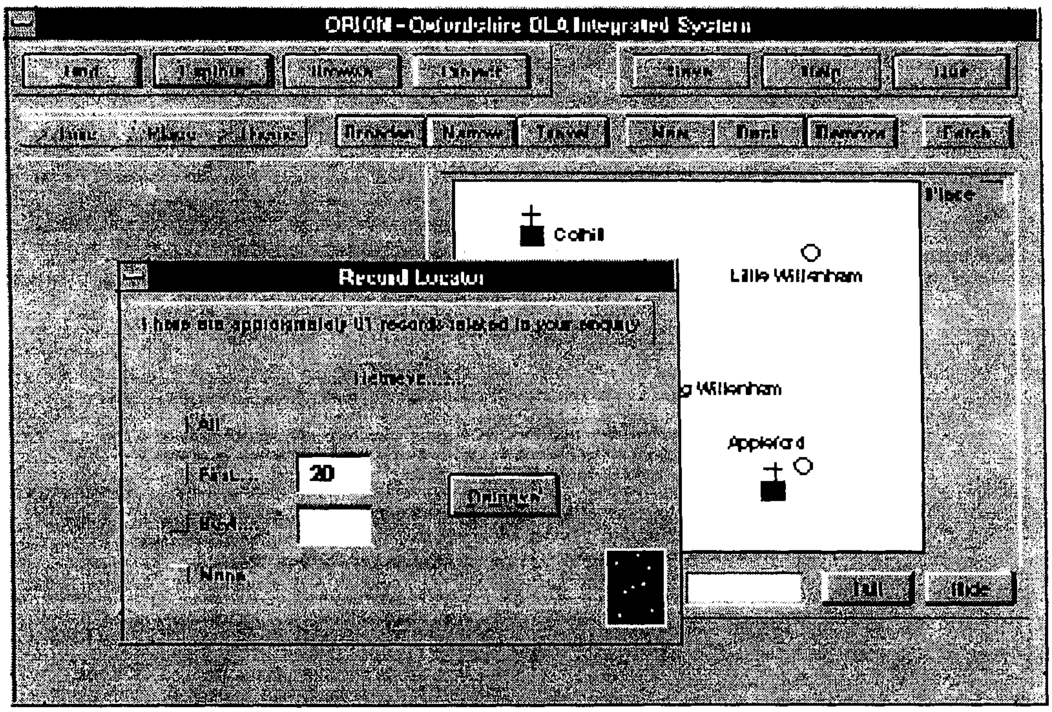

Figure 14

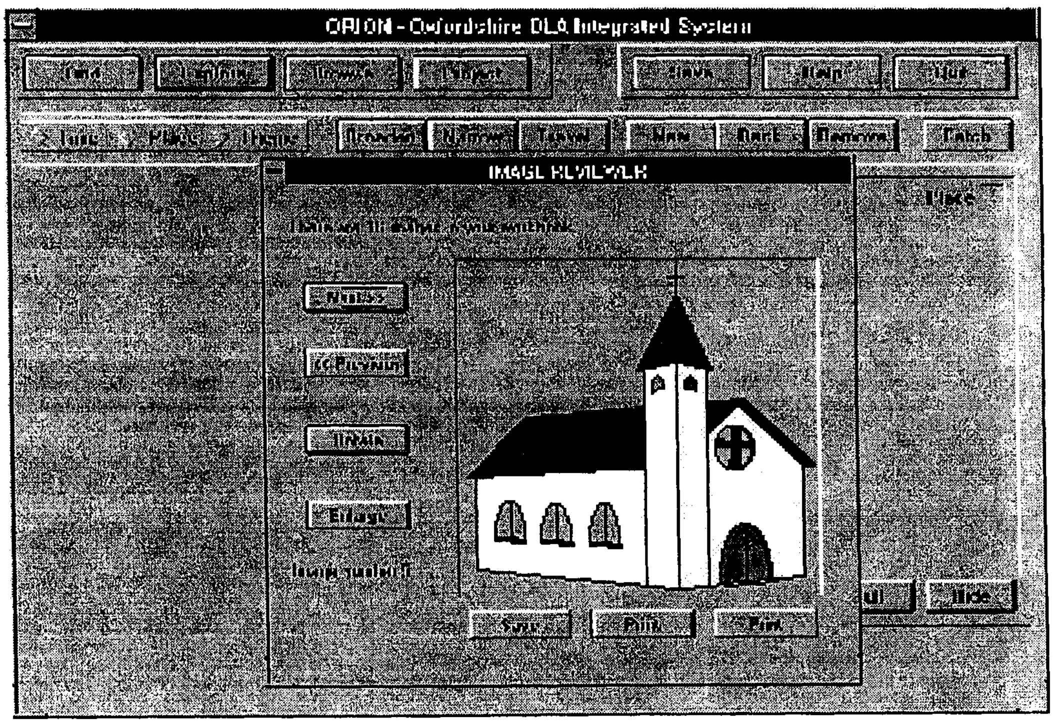

Figure 15

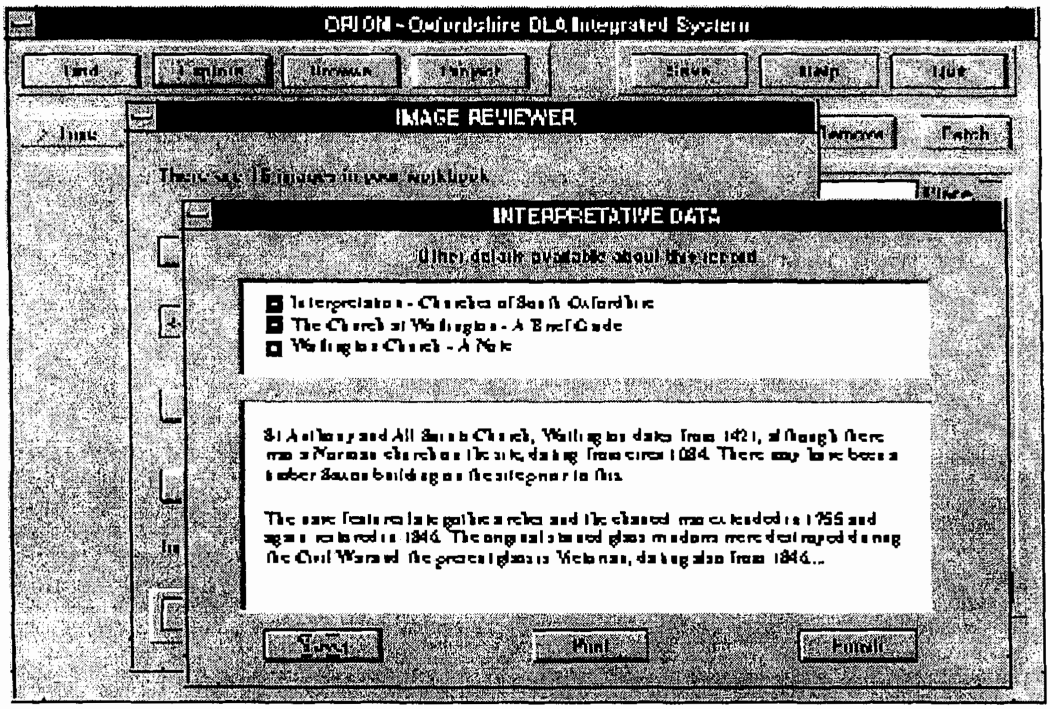

Figure 16

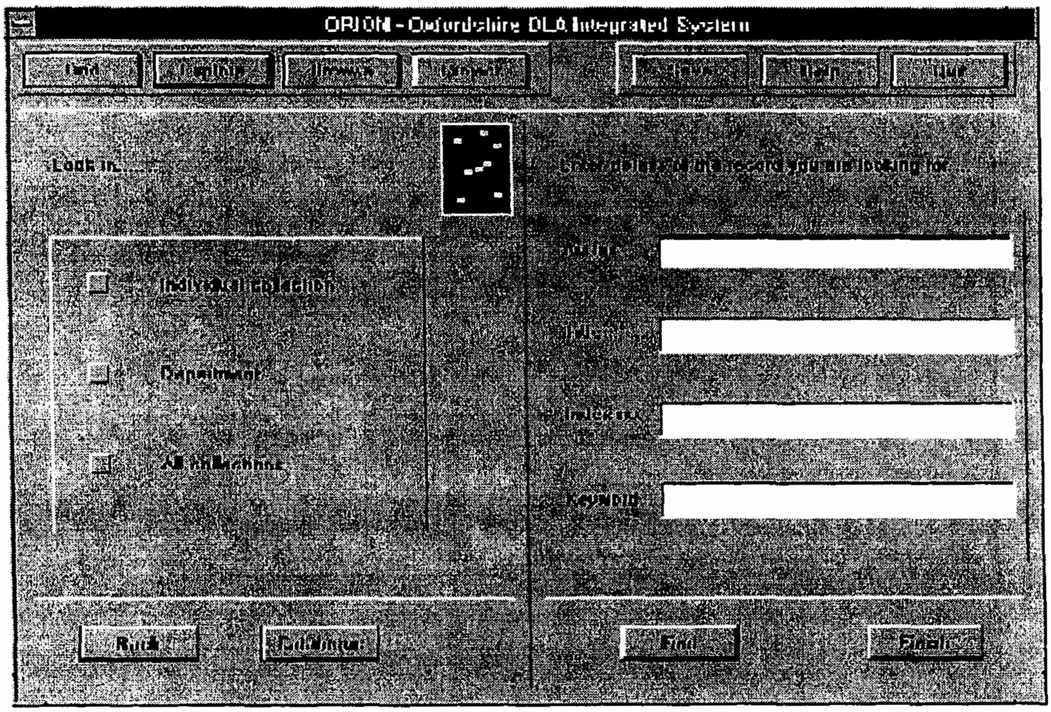

Figure 17

 © Archives & Museum Informatics, 1997

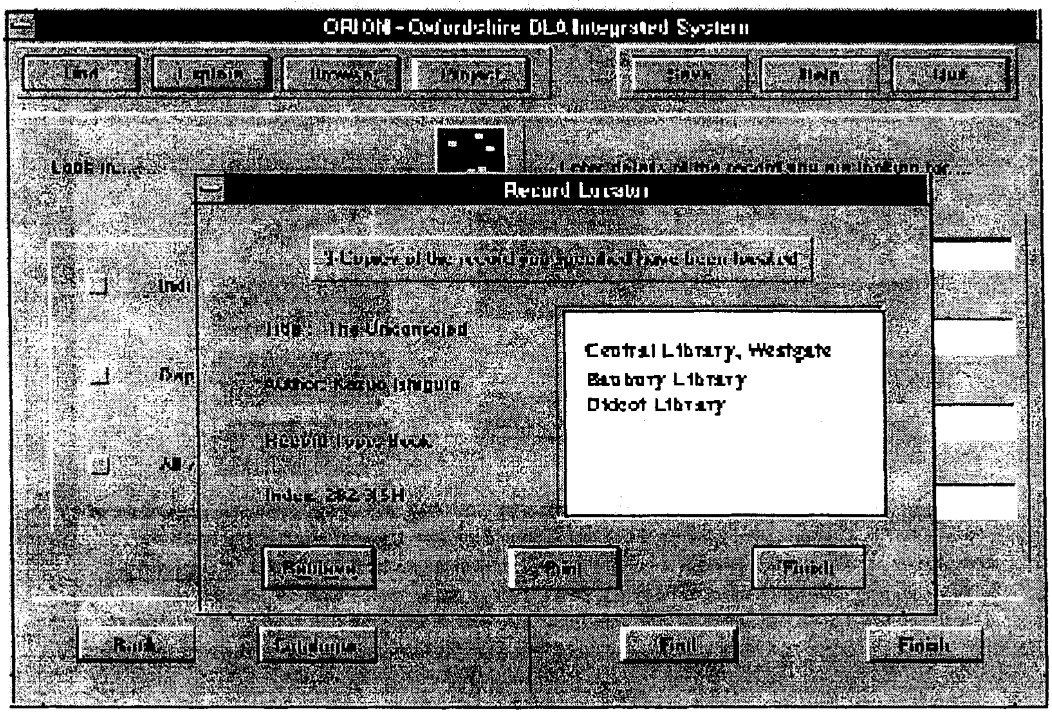

Figure 18

propriate in the cases of items where such representations exist; in other cases the Retrieve" button will be inactivated.

THE BROWSER

The third element of the ORION interface serves the needs of users who simply want to browse through what the DLA has to offer. To this end, a modified Internet browser is supplied, illustrated in Fig 18. Users simply traverse hypertext links in the conventional way; the interface has been deliberately simplified from that offered by, say, Netscape.

The ORION specification offers a number of other facilities, which space forbids a detailed description of here, including a Helper—the visual manifestation of a software Agent known as the Semantic Agent. The Helper provides on-line, context-sensitive help facilities for users of any of the ORION interfaces. If the user presses the button marked "Help" at any stage. a small box will become active and can be floated to any position on the screen. There are a number of kinds of help that the Agent can supply. These are

1. Help with the meanings of terms. The system can provide information about what certain words thrown up by the system replies or instructions mean;

2. Help with how to use the system;

3. Help with finding items that the user is not able to specify in the Explorer, the Finder or the Browser.

All this help is *context sensitive*; that is, the system is able to detect what the user is doing or what areas he is concentrating on at any point, and only offers help in that area initially.

RETRIEVING RECORDS

DESIGN CONCEPTS

A full discussion of the computational methods used for the actual retrieval of records and representations across the collections of the entire County, forming as it does the longest and by far the most technical component of the specification, would be out of place here. The purpose of this section is to

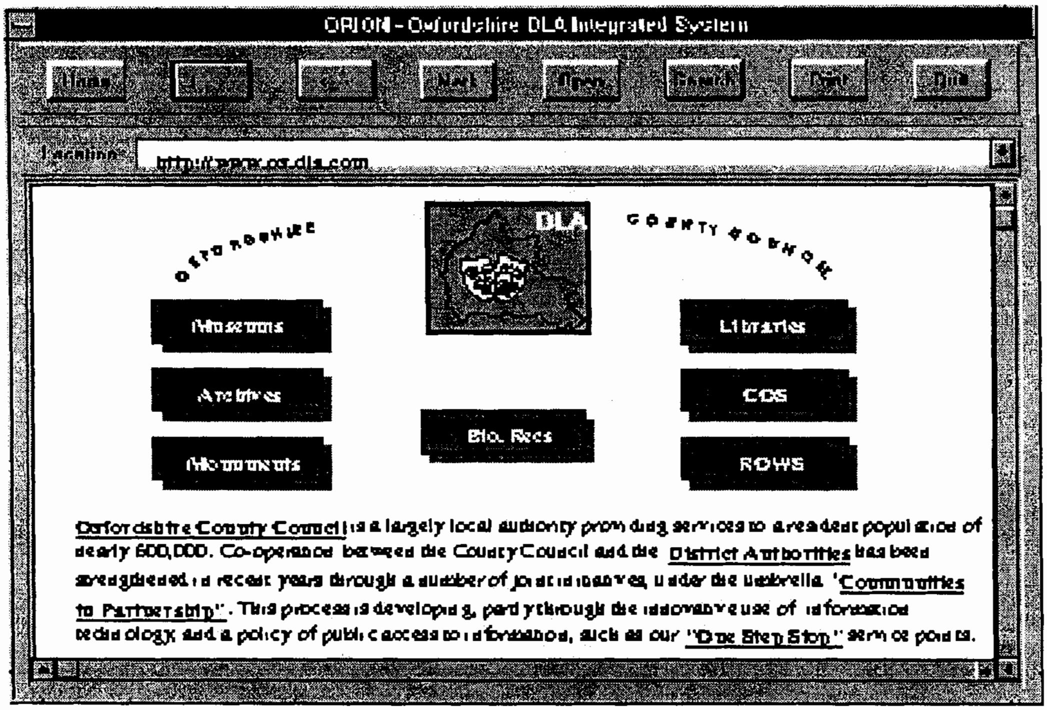

Figure 19

give an outline discussion of the principal design concepts and components of the ORION software.

Five themes run through the design of the system and appear in the *O*utline *S*ystem *D*esign (OSD) of the ORION system. These are as follows:

- Graphical Interaction
- Space, Time and Theme
- Agents
- Multilevel Indexation
- Auto-translation

Each of these topics is now discussed in more detail.

Graphical Interaction

The systems through which users should interact with the system through a series of graphically-based point-and-click system has been described in the preceding sections; but our interface design carries through into the design of lower-level components of the system. These graphical elements have a direct mapping onto the concepts that the user is investigating; and onto the sub-systems that co-operate with users in their searches and browses;

Space, Time and Theme

These concepts also pervade the design of the ORION system, manifesting themselves in numerous low-level software components;

Agents

Agents are autonomous software sub-systems with specialized capabilities that collaborate to fulfill some user requirement. Users interact directly with certain Agents at the HCI level—a specific Agent may be responsible for handling some Theme, for instance; other Agents, with capabilities of indexation, or translation, for example, are co-opted dynamically in the process of solving some problem. Agent technology has a reputation for being at the forefront of research into the design of intelligent software systems [Pfeifer & Verschuren, 1995]; but we have chosen it for a much more serious reason—we believe that a system composed of a number of discrete, collaborating sub-systems, with specialized

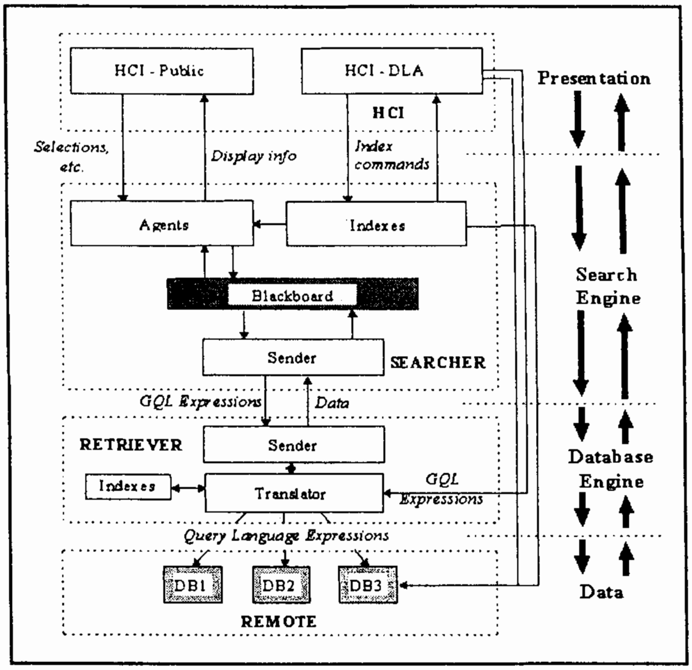

Figure 20

capabilities is the best way of designing and implementing systems that are robust, reliable and straightforward to test and upgrade as requirements change.

MULTI-LEVEL INDEXATION

Individual items in collections are accessed through a set of linked and layered indexes. These indexes are used by the appropriate software Agents to produce a profile of the user's current information requirements, which are then expressed in an internal language called General Query Language (GQL), based on work on general purpose languages for interfacing to distributed databases [Cohen & Ringwood 1993]]

AUTO-TRANSLATION

GQL statements are handed from the Searching sub-system to a retrieval system that locates the appropriate external databases, and converts the GQL statements into a set of extraction statements in the query languages appropriate to each database. These commands are then passed to the external databases. GQL statements can also be entered directly into the Translation system through the DLA's internal interface.

IMPLEMENTATION CONCEPTS

The authors of this document believe that there is only one way to analyze, design and implement software systems: Object-Orientation (OO) [Meyer 1997]. This is particularly true of a system such as ORION in which a number of independent agents collaborate, consulting multi-level indexes, in which there is a high degree of graphical interaction, and in which there is a strong requirement for configuration without re-programming. OO is now a mature system-construction technology, used by software developers throughout the world; it is fair to say it is probably now the technology of choice for developing any large scale software system. Its rationale is to produce systems that are robust, testable, allow upgrading and other change with minimal disruption to existing software components {Graham 1991, Eliens 1995]. A plethora of languages (C++, Smalltalk, Java, Eiffel, etc.) and design methodologies (HOOD, OOD, Booch, BON, etc.) have sprung up to provide support for the O-O approach. The physical design of ORION is entirely object-oriented, but will not be discussed further in this paper.

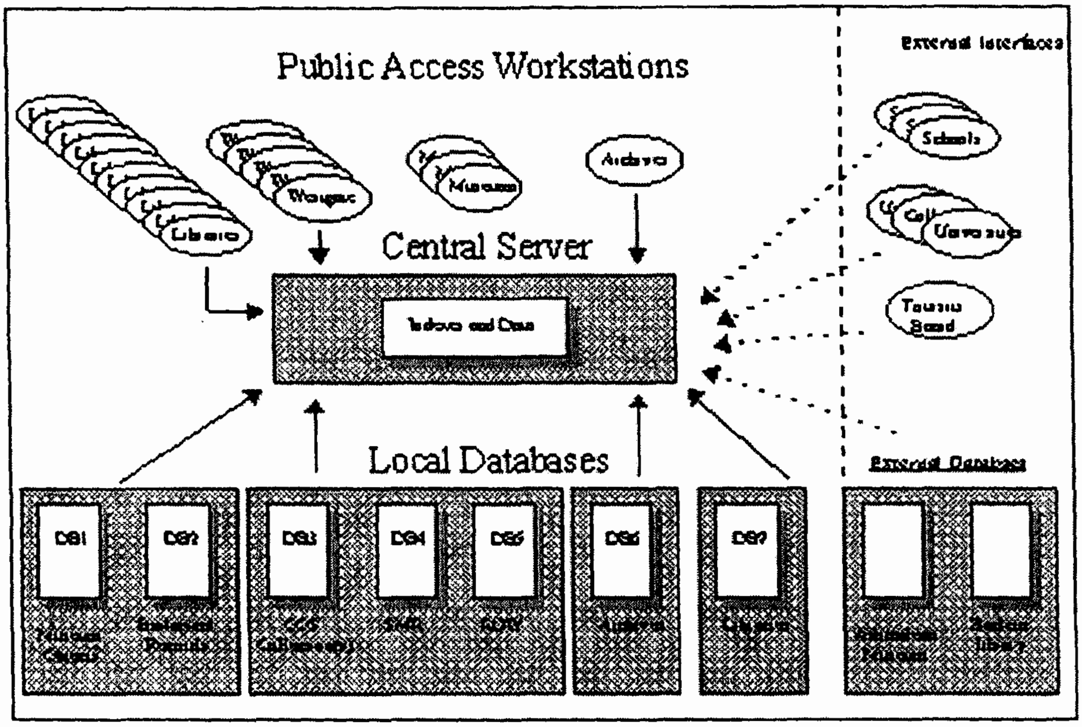

Figure 21

Outline System Architecture

Fig 20 expresses a high-level view of the proposed system's logical architecture:

Fundamentally, we propose a four-layer model, the four strata of the system being as follows

1. The HCI layer, contains all the software components that handle the interactions between the system and the various types of user. These interface elements were outlined in the preceding section.

2. The Searcher handles the processes of exploration, broadening, narrowing and traveling in the dimensions of time, space and theme that were described above—supplying data for the HCI to display graphically and communicating with sets of indexes to retrieve information about the number and type of DLA records within the current focus. These interactions are handled by specialized software agents

3. The Retriever receives messages from the Searcher coded in a common format, General Query Language (GQL). The purpose of the Translator is to locate the remote database at which a certain DLA record is held and to transform GQL expressions into expressions in the database query language employed by that database. The Translator also handles the interaction with that database.

4. The Local Databases are currently employed by the DLA in their museum catalogues, study centers, archives, etc. and are located at sites around the County; these supply data to the Retriever, which then passes is back to the HCI for display to the user.

The Distributed Database

Fig 20 illustrates the physical hardware configuration on which we are proposing that this software should run. In discussing its physical architecture, major questions arise as to how this architecture should be distributed over sites in Oxfordshire: the physical location of data; the physical locations of indexes; the physical location of the software subsystems we describe in this section; and the Client/

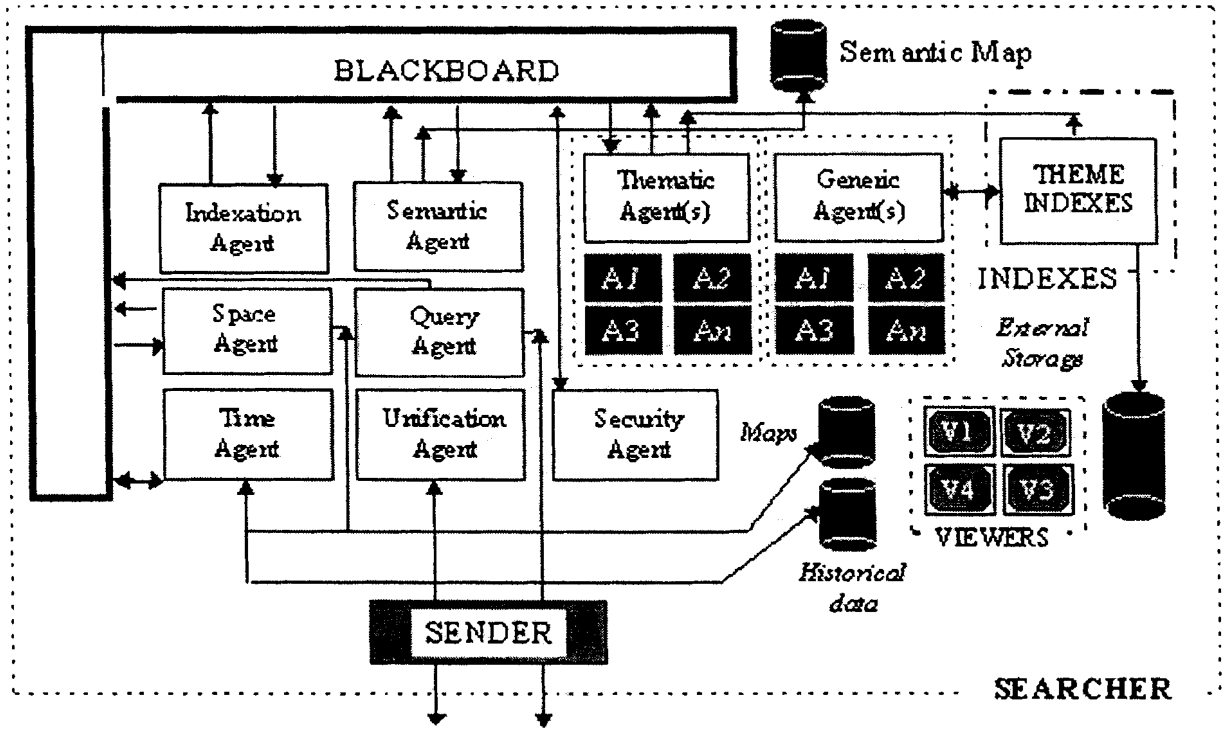

Figure 22

Server interactions implied by distribution choices. As Fig 21 shows, data may be located in individual machines in diverse locations around the County (e.g. Central Library, Archives, Local Libraries, etc.). The very important point, however, is that *any* data can be retrieved and passed to a public access workstation, *wherever* it is located. In ORION, we are proposing a *distributed database*. From the point of view of a user of the ORION system, it is irrelevant *where* data is stored: it can be located and returned without the user having to know any facts about location.

THE SEARCHER

An expanded view of the logical design of the Searcher software layer is represented in Fig 22 The main components of this layer, each of which is designed in detail in the specification, are as follows:

- A set of *Agents*. Each agent is a discrete, independent software component with responsibility for some aspect of the layer's overall function [Smithers 1995];
- A *Blackboard*. This is an area where agents post the results of operations they have completed and read the results that other agents have posted there: it is thus a means (the only means) of intercommunication between agents;
- A set of *Indexes*. This set consists of a sub-set of *Theme Indexes*, each responsible for a particular main Theme, and a *Names Index*, referencing specific items in the DLA collection; the indexes are retained by a combination of internal and external storage.
- Other external storage will include the *Maps* and *Historical Data* that the Space and Time Agents will require for display of the 3-d and 2-d views of the county.

The Searcher receives inputs from the HCI layer in the form of user selections, key-presses, mouse clicks, etc. and at which it makes displays, through the software Agents. It sends input to the next layer down, the Retriever, in the form of GQL statements, and receives confirmation and data back from the Retriever in the form of acknowledgments and data to be displayed at the HCI layer.

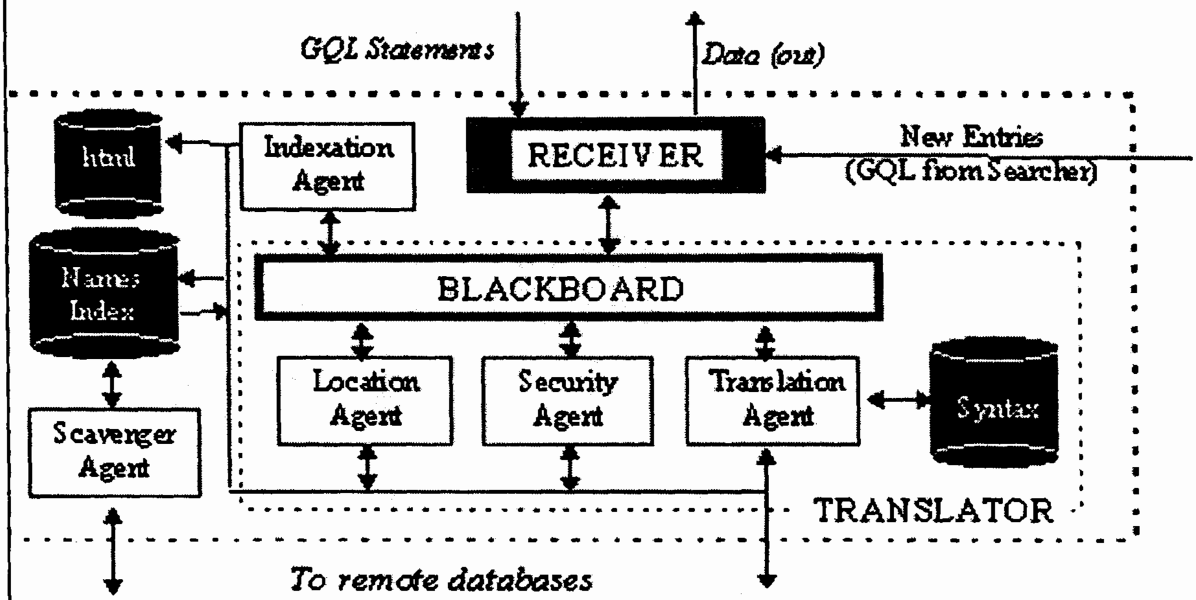

Figure 23

Each of the Searcher's Agents will fulfill a number of generic tasks:

- *Specific functionality*: each agent is responsible for certain concepts, functions, or dimensions. For instance, the Time Agent will be responsible for all functions to do with the Time dimension of exploration discussed above;
- *Interaction with the user*. Some—though not necessarily all—Agents have a visual manifestation at the HCI level: the Space Agent, for example, will be responsible for receiving user selections of some area of the county on the Space dimension and for managing display of 2-d and 3-d representations of that space;
- *Collaboration with other agents*: Obviously, agents will need to collaborate in order to solve problems, since each agent, with its specialized capabilities, can only work on part of a user's requirements. Consequently, there is a need for agents to share information—this is done by means of the blackboard;
- *Learning*: Agents have a limited capacity to act on the 'experience' they acquire during system operation over time. For instance, an agent for a particular Theme may store information showing that a particular breakdown of that Theme has frequently been called for, or that a particular user has, in the past, often exploited certain links or types of operation. This may allow the agent to perform proactive behavior, suggesting certain actions, before they are actually called for. A further benefit of such stored experience is that DLA managers of the system will have records of how ORION is used by the public, enabling them to configure it in effective ways.

The Retriever

The Searcher layer described in the previous section acts broadly as a means by which users may specify their information needs, touring a virtual representations of the county in space and time. The Retriever acts as an integrator, accepting the client's specification of their needs, searching all DLA catalogues, presenting clients with lists of records that lie within their current Focus, and locating and returning selected items in the DLA collections for a user to inspect. The Retriever is aware of such issues as of copyright, confidentiality, security, etc. and will enforce the regulations by denying access, where applicable. The logical structure of the Retriever is given in Fig 22:

The main components of this layer are as follows:

- A set of *Agents*. The same Agent technology is used as in the Searcher. None of the Agents of the Retriever have any visual manifestation at the HCI.
- A *Blackboard*. This has the same function as in the Retriever, that is as a means of communication between agents;
- The *Names Index*. This references specific items in the DLA collection; the indexes are retained by a combination of internal and external storage.
- A *Syntactic Database*, containing details of the schema's of the all local databases—and the query languages used to access them, thus allowing the Translator to formulate appropriate queries on remote databases, in cases where information is not stored on the central server;
- *HTML pages* that can be developed separately from other Retriever software and accessed by conventional Internet means, mediated by the Indexation Agent.

The Retriever receives input in the form of GQL from the Searcher, and acts as an integrator across all the local databases situated at DLA sites across the county, retrieving data from the Names Index and from these local Data Stores, as appropriate and returning it to the Searcher to be displayed at the HCI.

This completes our survey of the ORION system architecture. The technical details of the schemata of the various databases and their indexes, the internal structures of agents, the syntax of the GQL, etc. are not discussed in this paper: details can be found in [OBU 1997].

CONCLUSION - A LIVING SYSTEM

The specification of ORION was carried out in a comparatively short time, and on a comparatively low budget. A number of prototypes were produced as proof of concept, but clearly the full scale implementation of the system will be a very large-scale effort, which we have priced at approximately £6 million (including hardware and Intranet capabilities). Major funding is required and the technical specification has been offered as a bid for funding through the UK Heritage Lottery fund; a consortium including Oxford Brookes University, the Universities of Santander and Antwerp, the Oxfordshire County Council Department of Leisure and Arts, and similar bodies in other European cities, has been formed to bid for European Commission Telematics funding. An early continuation of the project, providing funding can be found, would be the building of a prototype of the user interface described above for intensive usability testing on members of the general public; a more detailed project plan can be found in [OBU 1997].

When completed and running, we believe ORION will serve a wide community of users, from the most rigorously scholarly, to the most casual. Moreover, we see it as a *living system*: as new items are added to collections, as new sites are discovered, they may easily incorporated into ORION; new interpretations of existing items may be added, as scholarship advances; the system may be extended to take advantage of new multi-media technologies. We hope and believe it will be an asset to all the people of Oxfordshire and its many visitors, imparting to all who use it a true sense of place.

REFERENCES

Bertin, J., *Graphics and graphic information processing* (W de G, 1981).

Brown, C., *Human-computer Interface Guidelines* (Ablex Publishing, 1988).

Carroll, J., *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction* (MIT Press, 1987).

Cohen, D. and G. Ringwood, "Distributed Databases Tied with StrIng" *BNCOD 11* (Springer-Verlag, 1993): 76 - 92.

Day, D. and D. Kovacs (eds.), *Computer Communications and Mental Models* (Taylor & Frances, 1996).

Eliens, J., *Object-Oriented Software Development* (Addison-Wesley, 1995).

Graham, I., *Object-Oriented Methods* (Addison-Wesley, 1991).

Meyer, B. *Object-Oriented Software Construction (2nd Edition)* (Prentice Hall, 1997).

OBU, ORION Technical Specification (1997).

Monk, A. (ed.), *Fundamentals of Human Computer Interaction* (Academic Press, 1984).

Microsoft, *The Windows Interface: An Application Design Guide* (Microsoft Press, 1992).

Pfeifer, R. and P. Verschuren, "The Challenge of Autonomous Agents: Pitfalls and How to Avoid Them" in L. Steels and R. Brooks (eds.) *The Artificial Life Route to Artificial Intelligence* (1995).

Smithers, T., "Are Autonomous Agents Information Processing Systems?" in L. Steels and R. Brooks (eds.) *The Artificial Life Route to Artificial Intelligence* (1995).

Schneiderman, B., *Designing the Human Computer Interface* (Addison-Wesley, 1992).

© ARCHIVES & MUSEUM INFORMATICS, 1997