CULTURAL DOCUMENTATION: THE CLIO SYSTEM

Panos Constantopoulos

University of Crete
and
Foundation of Research and Technology - Hellas

Institute of Computer Science
Foundation of Research and Technology - Hellas
P.O. Box 1385, 711 10 Heraklion, Crete, Greece

CULTURAL DOCUMENTATION: THE CLIO SYSTEM

Panos Constantopoulos

University of Crete
and
Foundation of Research and Technology - Hellas

Abstract

Cultural documentation, i.e. the recording and management of a body of knowledge about ensembles of cultural goods, presents special requirements of a computer-based information system, not met by ordinary documentation systems and cultural information bases. The CLIO system aims at covering these needs, while it can also cooperate with an administrative documentation system. Information is organized in CLIO as a knowledge base according to a specifically designed semantic model. The construction of CLIO allows extremely dense linking of information, access by unlimited chained references, expression of abstract properties and various ways of joint temporal and spatial assignment. It also allows the extension and modification of the data schema by the users, thus supporting the easy adaptation of the system to the field of work and the evolution of knowledge.

1. The documentation of cultural goods and the CLIO system.

The documentation requirements of cultural goods range from keeping a simple log of objects to the management and conservation of collections and even to the recording of a variety of cultural information. We distinguish documentation into administrative, using a fixed set of data mainly to support administrative functions and to provide basic information, and cultural, aiming at organizing an evolving body of knowledge about objects, to be used in scientific study and research. The existence and interconnection of both is ideally desired.

Most computer-based systems support administrative documentation and we shall call them administrative documentation systems even if they are mere log systems. They usually employ relational data base management systems, ideally suited for organizing information in table format, value-based search and report generation. On the other hand, the relatively recent multimedia information systems are mostly "closed" systems for providing multimedia information addressed to the public, featuring fixed sets of information and navigation paths through it.

Cultural documentation requires a system capable of recording the entire variety of information, which constitutes the current knowledge about a set of objects. This includes formatted data and other, multimedia data (images, audio and video recordings, text...) and are characterized by a high degree of linking, large variety of references and classifications, broad usage of abstract relations and the need for multiple, mainly referential access. Moreover, this information is in general enriched but rarely modified. Ordinary administrative documentation systems and hypermedia information bases do not meet all these requirements.

The CLIO system was developed at the Institute of Computer Science, Foundation of Research and Technology - Hellas to fulfill the requirements of cultural documentation. Information is organized in CLIO as a knowledge base according to a specifically designed semantic model. The functional kernel of CLIO is the Semantic Index System (SIS, see [1]), built at the Institute of Computer Science. The construction of CLIO allows extremely dense linking of information, access by unlimited chained references, expression of historical and cultural context as well as of abstract properties, joint temporal and spatial assignment in absolute or relative terms, and recording alternative, possibly conflicting information along with the respective sources. Information is presented in graphical or textual form. An extensible list of predefined queries is offered. A particularly important feature is the uniform treatment of schema and data, enabling the immediate extension and modification of the schema by the users themselves. CLIO can be characterized as "open" system, in the sense that it involves a variable system of knowledge chiefly addressed to the museum curator or the researcher, and from which closed information bases for the public can be generated. Finally, it offers an interface to an administrative documentation system.

The analysis of requirements for the CLIO system has been performed in close cooperation with the Benaki Museum and the Historical Museum of Crete. The development of the system was partly funded by the STRIDE and ESPRIT programmes.

2. Recording knowledge in the CLIO system.

The formatted documentation data are structured according to a model which renders an appropriate ontology and is expressed in a particular knowledge representation language.

2.1. Knowledge representation.

The structural part of the Telos knowledge representation language [2] is used for representing knowledge in CLIO. This offers the general mechanisms of attribution, classification and generalization, common in all the so called semantic, conceptual or object-oriented representation methods, and, in addition, it supports unbounded classification hierarchies and equal treatment of relations and entities, resulting in great expressiveness and flexibility.

Attribution is a general mechanism for representing the properties of an object as relations between the object and the value ranges of the properties, also considered as objects. Classification defines classes of objects on the basis of common properties. An atomic object is declared as instance (member) of one or more classes. Classes are themselves considered as objects (sets), instances of meta-classes (sets of sets), and so on. Thus, levels of classification are distinguished: tokens, classes, meta-classes, etc. Generalization defines a subclass-superclass relation between classes at the same level of classification, characterized by the inheritance of properties: when class A is subclass of B, it inherits all the properties of B and is differentiated from it either by having additional properties of its own, or by restricting the value ranges of the inherited ones. Finally, every object (entity or relation) has a unique identifier.

2.2. Ontology of CLIO.

The knowledge representation model of CLIO expresses a rather general ontology which can be complemented by derived concepts, thus specialized to particular fields. This ontology includes concepts of matter, location, occurrence, quantity, mankind, spiritual creation and naming, as well as relations between them, and is presented in detail in [3]. Here, we are restricted to a very brief review.

The concepts of matter account for the kinds and the structure of physical objects, the elements of their appearance, the distinction into natural and artificial, composition or construction materials, special information concerning fine art objects and museum objects, descriptions of style and ornamentation, and kinds of tools.

The concepts of location define absolute and relative location, orientation and various topological relations.

The concepts of occurrence introduce certain fundamental notions, such as event, as a point in time and space, and existence, as a set of events delimited by a beginning and an end, and include categories of events of particular significance and such concepts as activity, use, creation, technique, etc.

The concepts of quantity distinguish between physical, arithmetic and numismatic quantities, elaborate their kinds, and define ways of measurement. The latter are distinguished on one hand into absolute, defined with respect to conventional coordinate systems (e.g., chronology, location), and relative, expressing magnitudes in given measurement systems (e.g., time, distance, area, volume, weight), and on the other into exact, yielding a single value, and approximate, yielding a range of values.

The concepts of mankind concern the description of human persons and of manifestations of activity and organization, such as subject, group, membership, institution.

The concepts of spiritual creation essentially differentiate the product of spiritual creation from its physical embodiments.

Finally, a system for naming and identification is defined.

The CLIO knowldge representation model subsumes the CIDOC/ICOM fine arts documentation standard.

2.3. Using the abstraction mechanisms.

The chief abstraction mechanisms, classification and generalization are extensively used in CLIO for reasons of economy and, mainly, expressive power. Classification, most notably multiple, is recommended for expressing intrinsic properties. The usage of higher levels of classification enables the expression of abstract or general properties which are of interest as such even when instantiated differently in particular cases (e.g., partition, overlapping, creation, measurement interval). The usage of multiple generalization enables the creation of entangled hierarchies, suitable for representing complex terminological systems and faceted classification schemes, in addition to great economy of expression.

3. Structure and function of the CLIO system.

The structure of the CLIO system together with its functional incorporation in an integrated museum information system are shown in figure 1. Such an integrated system offers combined administrative and cultural documentation functionality by interfacing respective autonomous systems. Administrative and multimedia data are acquired and managed by the administrative documentation system, using a relational data base management system to store them. Part of the administrative data is also transferred to the cultural documentation system for purposes of object identification and matching certain basic information. Furthermore, multimedia data are retrieved from the common store for presentation by the cultural documentation system. Details about the architecture of the integrated museum information system, designed at the Institute of Computer Science, FORTH, and of CLIO, as part of it, can be found in [4]. In the sequel the structure and function of CLIO are briefly presented.

CLIO consists of three subsystems: the semantic index (knowledge base), the user interface and the communication interface. The semantic index stores and manages information organized according to the knowledge representation model (see 2.1), using the Semantic Index System (SIS, see [1]). The communication interface comprises the data transport system, performing batch import of selected data from the administrative to the cultural documentation system, and the retrieval of multimedia data. Finally, the user interface offers the functions of the system to the user, thus determining its perceived behaviour. The user interface supports extensive adaptation of usage dialogue and specialization of functions to meet special requirements of the field of work, without reprogramming.

The functions of the user interface are distributed over three subsystems: data entry, information retrieval and browsing, and multimedia display. Most important, from a functional point of view, are the first two.

Data entry is carried out through a special form which offers guidance and vocabulary support and control. The data entry form is dynamic, in that it adapts to the current task automatically. The data entry function uniformly addresses both the entry of data conforming to a given schema and the modification of schema information. Thus, the extension or correction of an information structure is a simple job, subject, of course, to authorization constraints.

The information retrieval and browsing subsystem presents segments (views) of the knowledge base in graphical or textual form (see figures 2 and 3). Browsing takes an item of the knowledge base, called focal object, as point of reference. Queries may concern the objects related to the focal object through specific relation types (1st order queries), or they may express recursive searches generating chains of objects and relations rooted at the focal object (recursive queries), or they may address combinations of attribute- or classification- based predicates. General or frequent queries are offered in lists of predefined queries for ease of use. These lists are easily constructed and modified, adapting to the terminology and requirements of the field of work. Predefined queries can be abstract expressions, the semantics of which may vary according to the situation (e.g., "main characteristics", see figures 3 and 4). Altering the focal object results in navigation through the mesh of references which constitutes the knowledge base. Referential access to information is deemed more natural in a system which, essentially, manages a large variety of relations. Finally, a session history is kept to prevent disorientation and facilitate the repetition of previous transactions.

The operating environment of CLIO comprises UNIX system V or SUNOS or HP-UX 8.05, X11 Rel. 4.5 window system and OSF Motif 1.1. The use of a colour monitor is to be preferred.

4. Applications.

CLIO can be used for the documentation of families or collections of objects, on condition of prior enhancement with specific field knowledge, achieved through the data entry subsystem.

In the context of testing and demonstrating the system, five prototype applications have been developed: a catalogue of cretan and byzantine pots, a catalogue of greek folk costumes, a catalogue of a collection of religious items, a catalogue of a collection of weapons and a catalogue of a painting exhibition. The catalogue of cretan and byzantine pots was originally developed as a student project with the aid of collaborators of the Historical Museum of Crete, subsequently carried over to CLIO (figures 2 and 3). The catalogue of greek folk costumes is developed in cooperation with the Historical Museum of Crete and will be used in the documentation of its respective collection. The catalogue of religious items and the catalogue of a collection of weapons concerns collections of the Benaki Museum and is developed in cooperation with the Documentation Department of the museum (figure 4). Finally, the catalogue of a painting exhibition is the outcome of a student project and concerns the exhibition "From El Greco to Cezanne" which took place at the National Gallery in 1993. In this project the CLIO system was employed, yet the user interface was replaced by another, tailored to the documentation of paintings and addressing a larger public (figure 5).

Acknowledgements

The CLIO system is the outcome of a collective effort. Among my colleagues in the Information Systems and Software Technology Group at the Institute of Computer Science, FORTH, thanks are especially due to Martin Doerr and Maria Christoforaki for their contribution, and also to Chryssoula Bekiari and Maria Theodoridou. Costis Dallas, Eleni Tsenoglou and Iphigenia Dionyssiadou provided precious help, being our main external collaborators. Finally, thanks go to Manos Theodorakis and Yannis Tzitzikas for getting involved in the application of CLIO beyond the call of duty.

References

[1] P. Constantopoulos and M. Doerr. The Semantic Index System - A brief presentation. Institute of Computer Science, FORTH, May 1993.

[2] J. Mylopoulos, A. Borgida, M. Jarke and M. Koubarakis. Telos: Representing Knowledge about Information Systems. ACM Trans. on Information Systems, 8/4, Oct. 1990, pp. 352-362.

[3] M. Christoforaki, P. Constantopoulos and M. Doerr. CLIO: An Object-Oriented Model of Cultural Data. Report ICS-FORTH.MUIS.92.1. Institute of Computer Science, FORTH, Nov. 1992.

[4] M. Theodoridou, P. Constantopoulos, C. Bekiari and M. Doerr. Architecture of the Museum Information System. Report ICS-FORTH.MUIS.92.2. Institute of Computer Science, FORTH, Nov. 1992 (in Greek).

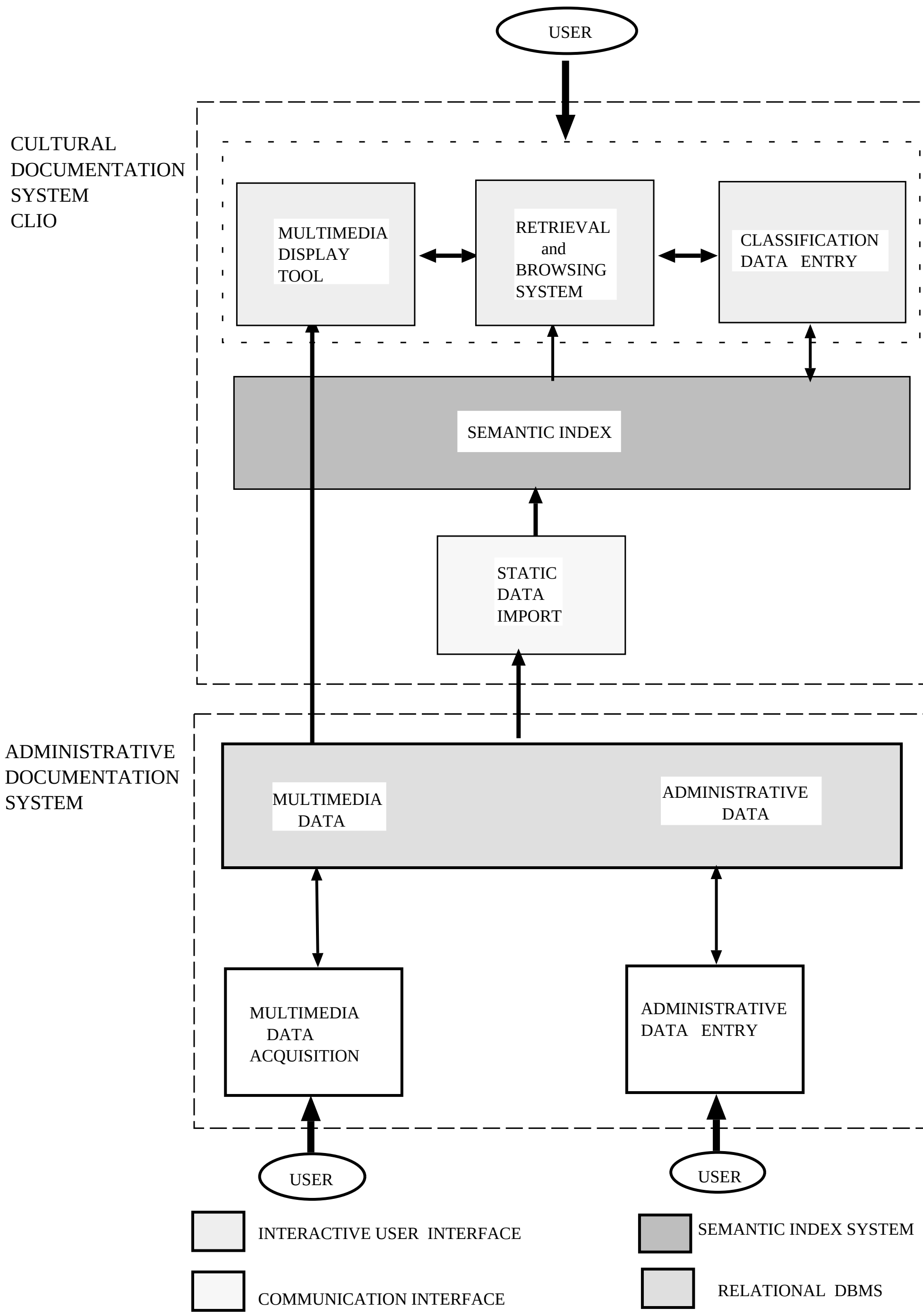

Figure 1: Integrated Museum Information System including CLIO

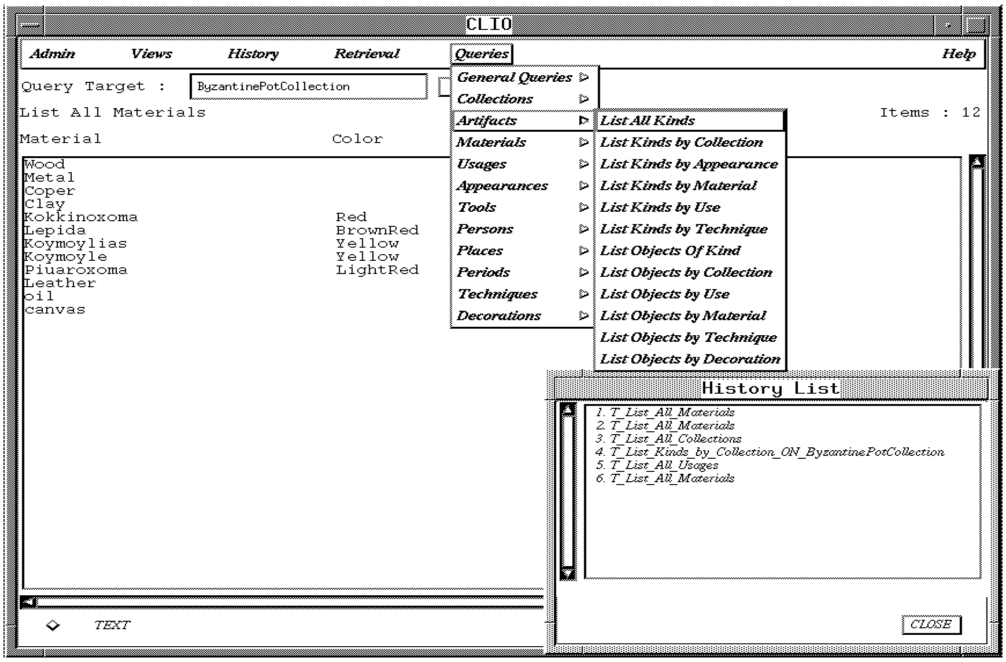

Fig 2a. 1st order predefined queries and session history.

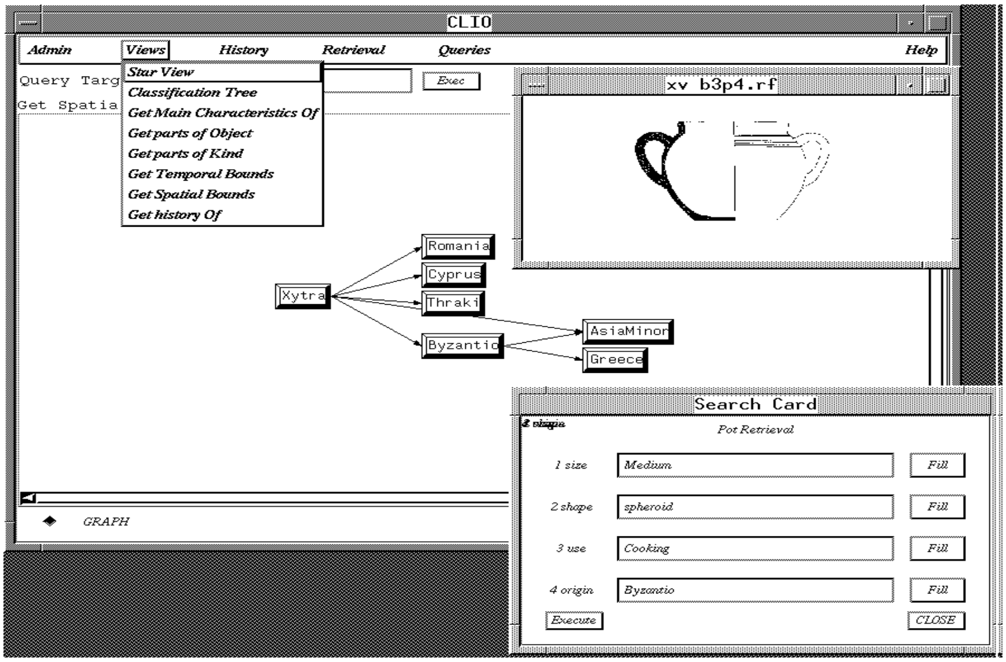

Fig 2b. Graphical view and search card.

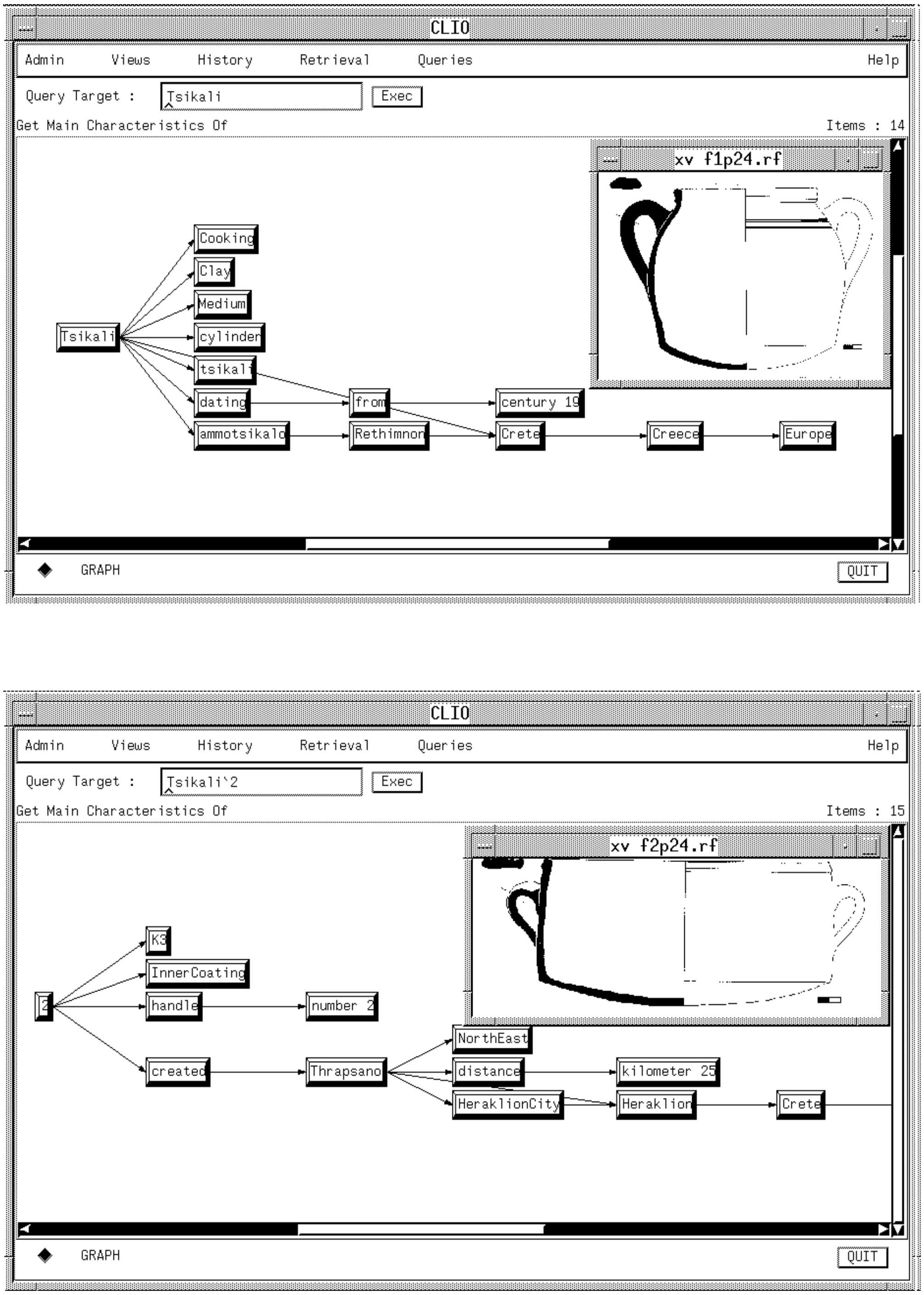

Fig 3. Recursive predefinied query about kind and specific object.

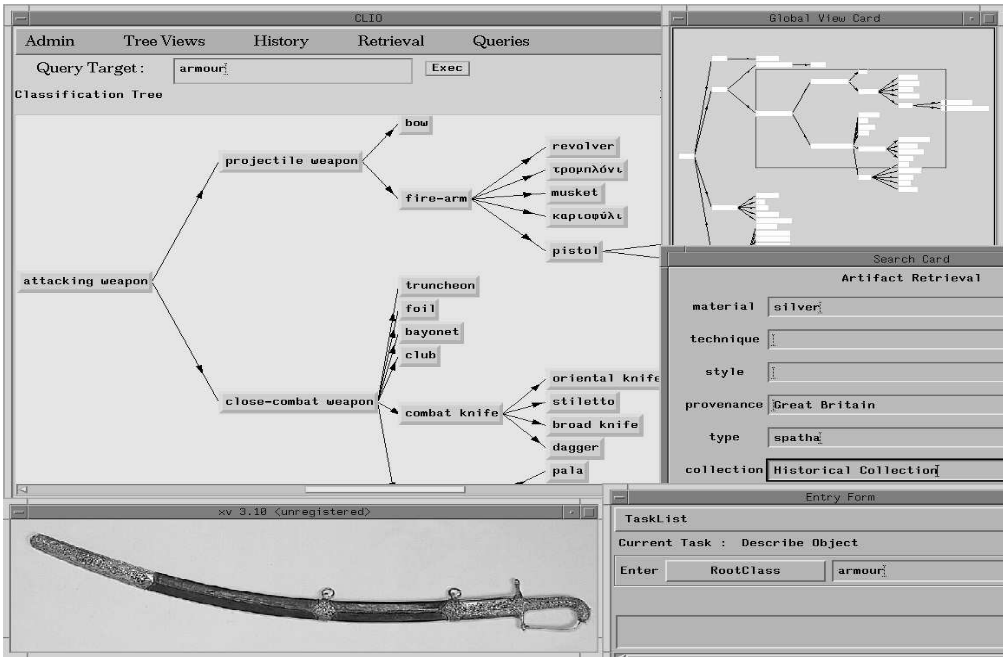

Fig. 4. CLIO: Weapon classification. Artifacts from the Historical collection of Benaki Museum

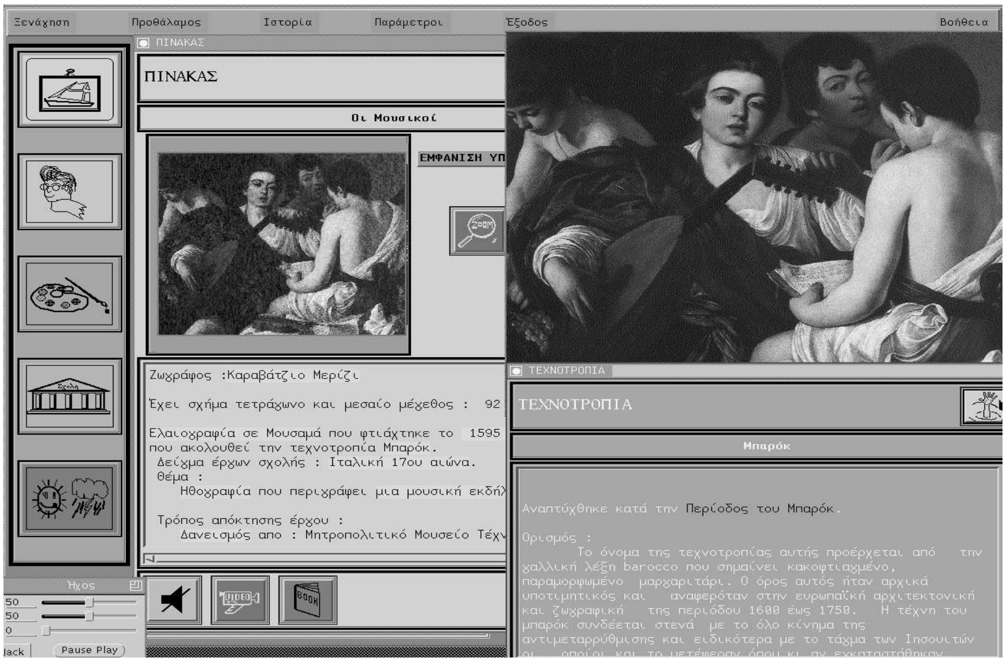

Fig. 5. Electronic catalog of painting exhibition: "From El Greco to Cezanne",National gallery, 1993 (Student project)