Hesperia: A Multidisciplinary Group-oriented Information System for Paleoclimatology

M. Vaquero, R. Theron

Departamento de Informática y Automática
Universidad de Salamanca, 37008
Salamanca, Spain
mvaquero@usal.es, theron@usal.es

J.A. Flores, F.J: Sierro

Departamento de Geología
Universidad de Salamanca, 37008
Salamanca, Spain
flores@usal.es, sierro@usal.es

Abstract— **HESPERIA is an information system for helping in the collaborative research of a multidisciplinary group in the paleoclimatology field. While there exist international databases and information systems that provide long term storage of georeferenced data, they cannot be used during the research stage, since they are limited to published data. The system is designed to manage a collaborative project, keeping track of the shared seabed samples, providing visual analysis and giving the possibility of using partial results. The basic technical structure corresponds to a three tiered client/server architecture that assures scalability and eases long-term maintenance requirements.**

Keywords: information systems, paleoclimatology

I. INTRODUCTION

One of the key obstacles of research institutions in the Paleoclimatology community is how to make available for the researchers all the involved data coming from a wide variety of proxies such us: forams, coccolithophores, diatoms, dynoflagelates, pollem (Micropaleontology), UK37 (Biomarkers), ^{14}C, δ^{18}O, δ^{13}C (Biogeochemestry), Paleoredox index (Geochemestry), Ba, Barite, Ti, Al (Sedimentology) and Granulometry.

Computational resources are helping to enhance the diffusion of this huge volume of data, allowing simultaneous access, efficient selection of information, analysis performance, and data and results sharing results among different interdisciplinary groups.

Although big efforts have been done in the development of international databases gathering relevant data for the field (i.e.: World Data Center for Paleoclimatology in U.S.A or Pangea in Europe [1]), they are limited to published data and give little opportunity for information and pre-results exchange for a particular group during the research process.

Systems that provide access to published data relevant for a particular paleoclimatic research project are of great value. Nevertheless, tools are needed that provide similar information but for not yet published data that can be shared among a multidisciplinary group. Scientists currently achieve this by sending e-mails to each other, usually with EXCEL files attached containing the needed data. Thus, it is difficult for the whole group to be aware of the current status of the project,

and some intermediate results from some members that could be used by some others members are not available.

Another important feature that a multidisciplinary (multi-laboratory and/or multi-institution) collaborating group can benefit from is the use of visualization and analysis tools for the shared data.

On the other hand, from the same core obtained through different ocean drilling techniques, several samples will be produced. Sometimes the number of potential experiments involved in a project is far more than the number of those that can be done due to sample shortages. Thus, it is very convenient for a group to keep track of the sample stock, the use of samples (in order to reuse when possible), and so on.

All these services can be provided within the same information system. This is the goal of HESPERIA.

The rest of the paper is organized as follows: in section II, the main features and system architecture are explained; the third section is devoted to the description of the visualization and analysis tools included in HESPERIA; finally, the main conclusions and future developments are summarized.

Figure 1. Sample tracking in the HESPERIA system

II. HESPERIA DATA MANAGEMENT AND STRUCTURE

HESPERIA (Figure 1.) is a web-based information system for storage, management and publication of seabed samples, cores and related information from the Mediterranean Sea. The

database is populated with data obtained in the European PROMESS 1 project (PROfiles across MEditerranean Sedimentary Systems Part 1). The HESPERIA information system aims to provide support for research activities and to facilitate coordination of PROMESS group by contributing to a better management and dissemination of marine data.

One of its main features is the management of shared samples (from ocean drilled cores) available for the collaborating research groups. For instance, for the sample tracking the following procedure is used: sample requests are stored together with all the participant institutions and laboratories, including information such as the requested volume and the analysis to be performed. This way, the curator of the PROMESS project can decide on accepting or rejecting a particular request.

Figure 1. shows the current use and stock for a sample of leg 120, site 1100, hole A, core 1. Additionally, a sample request report is automatically produced every week is sent as a PDF file by e-mail to the curator (Figure 2.).

Figure 2. A sample request report is sent peridiocally to the curator

The HESPERIA database is based on a relational data model consisting of over 100 tables. Multiple services (similar to the sample tracking explained above) are provided through the web regarding several aspects of the ocean drilling process needed during different research stages.

A. System architecture

HESPERIA is based on a three-tiered client/sever architecture [2] (Figure 3. . Each of these layers make use of an open source technology. In the typical three-tier architecture, the first tier is persistent data storage. In the case of our system we have chosen MySQL as the database server. The second tier is the application layer (middleware). This is where all the rules regarding the stored data are coded, i.e, it is the bulk of the HESPERIA system. This second tier is composed of a web server, Apache, running a scripting language (PHP) providing database access.

The third tier is presentation, enabling the users to see the results of HESPERIA rules applied to the stored data. The presentation layer is clearly the browser-application/browser, rendering the output of the PHP scripts.

This basic architecture assures excellent system administration, security, performance, scalability and application reuse [3].
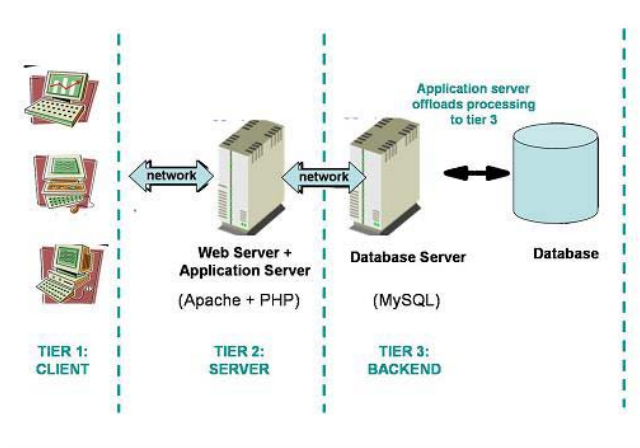

Figure 3. Three-tier Architecture

III. ANALYSIS AND VISUALIZATION

The main purpose of the HESPERIA information system is to be more than a web-based database. Thus, it was conceived as a research tool, including visualization and analysis features.

From the beginning, HESPERIA was designed to provide a full set of research capabilities while remaining the more easy to use for the users as possible. Thus, server-side technology was chosen so the user only would need a client browser.

One key aspect during the research process in the paleoclimatogy field is the creation of an age model. That is, to determine the age of the sediments from 0 meters to a particular depth of the core. This can be achieved using different techniques, and typically the scientist that provides an age model for a particular site will refine the model several times. This is very important since the rest of the researchers will work with ages instead of depths when modeling the past wind dynamics [4], for instance. Thus, both the creation and visualization of age models, and the possibility of plotting time series of paleorecords using different age models were specially useful tools that the HESPERIA information system should include.

Thanks to these tools, on-line tentative environmental reconstruction through real time analysis and visualization of data can be produced. Users simply upload their core data and age model files to the Hesperia server, or select the files of previous results of the group from the database.
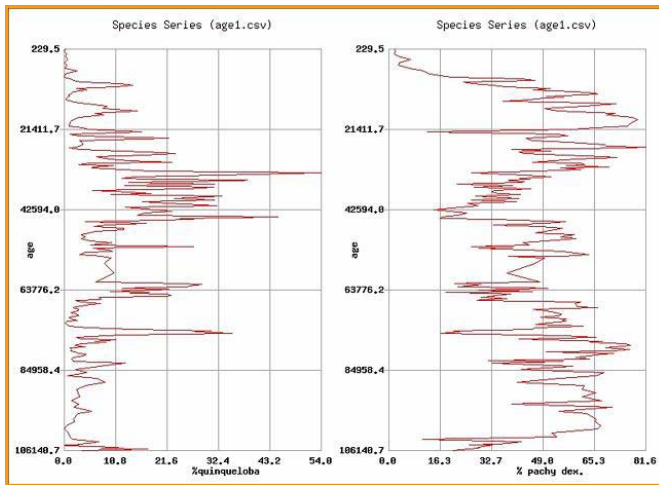
Figure 4. Several microfossil species paleorecords are plotted using a particular age model.

Figure 4. shows two microfossil species through time using a particular age model. This analysis tool permits the paleoclimatologists to interpret different dynamics in the past related to particular geological events. This process can be done any number of times using local data (either the species data or the age model) or data stored (either the species data or the age model) in the HESPERIA system.
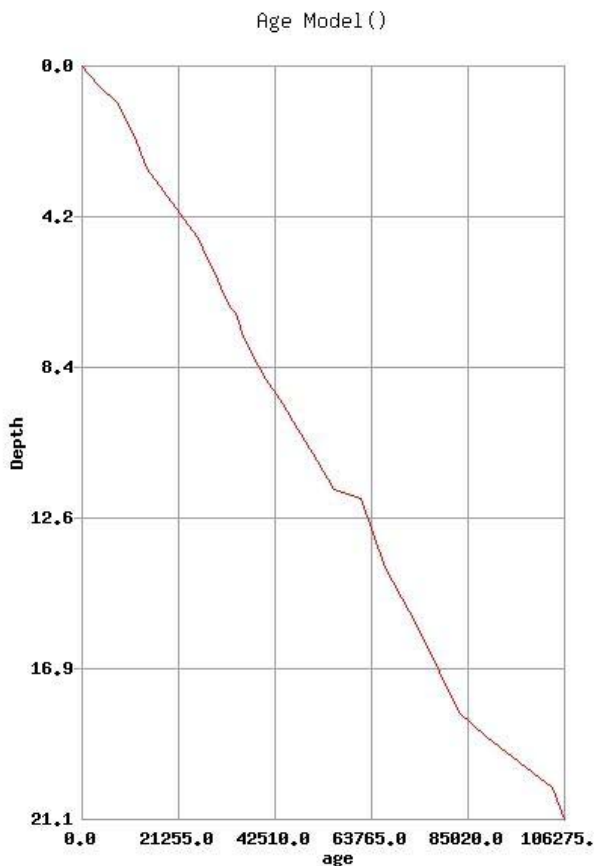
For each core an age model is generated based on several age index points under the assumption of a continuous sedimentary record. In order to provide the scientists with information about the age models that they are using, HESPERIA can plot the curves (Figure 5.).

Additionally, interval sedimentation rates are also plotted in order for the experts to better interpret the variations on the deposits by means of analysis of the most significant increases and decreases in the sedimentation rate (Figure 6.).
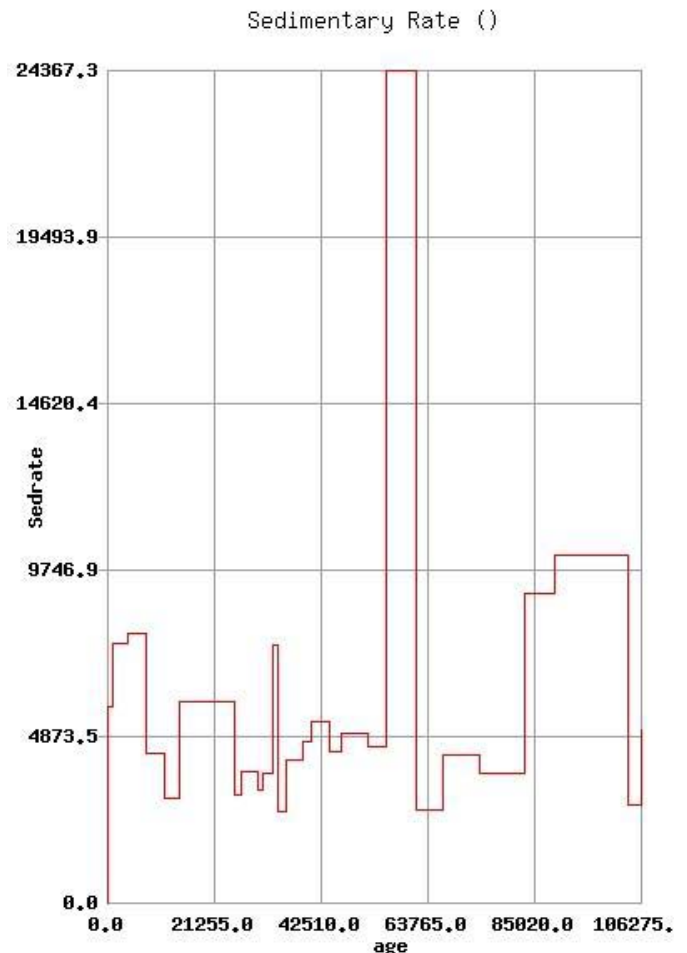

Figure 6. Sedimentation rate for a particular core

IV. CONCLUSIONS AND FUTURE WORK

HESPERIA is a very powerful research tool that fosters the collaboration among multidisciplinary groups from several institutions. Thanks to its three-tiered client/server architecture full simultaneous access from different locations is granted, so scientists can share data and partial results. Visualization and analysis of the datasets are provided such as age-model plotting and creation, as well as plotting of time series paleorecords using different age-models stored in the database. Currently, all the HESPERIA features are provided using server-side

Figure 5. Age model for a particular core

technologies. In the future some advanced tools for interactive analysis and visualization already developed by our group [5][6] will be included in the system as client-side tools.

REFERENCES

[1]   M. Diepenbroek, H. Grobe, M. Reinke, U. Schindler, R. Schlitzer, R. Sieger, G. Wefer, "PANGAEA—an information system for environmental sciences" in *Computers & Geosciences,* vol. 28, 1201-1210, 2002.

[2]   W. W. Eckerson, "Three tier client/server architecture: achiving sclability, performance and efficiency in client/server apllications". Open Information Systems 10, 1, 3 (20), 1—12, 1995.

[3]   S.S. Chiu, Y.P. Luh, "Architecture Design of Shop Floor Information System—Using WWW and 3-tier Architecture", International Journal of the Computer, the Internet and Management, vol. 9, number 2, may-august, 2001.

[4]   R. Theron, J. A. Flores, F. J. Sierro, C. Pelejero, J. Grimalt and M. Vaquero, "Using data mining and visualization techniques for the reconstruction of ocean paleodynamics," in proceedings of the *IEEE International Geoscience and Remote Sensing Symposium*, vol. IV, pp. 2382-2384, 2002.

[5]   R. Theron, J. A. Flores, F. J. Sierro, M. Vaquero and F. Barbero, "PaleoPlot: A tool for the Analysis, Integration and Manipulation of Time-series Paleorecords," in proceedings of the *IEEE International Geoscience and Remote Sensing Symposium*, vol. VI, pp. 3528-3530, 2002.

[6]   R. Theron, D. Paillard, E. Cortijo, J.A. Flores, M. Vaquero, F.J. Sierro, F.J., C. Waelbroeck, "Data-mining the past environment", in proceedings of the IEEE International Geoscience and Remote Sensing Symposium, vol. 6., pp. 3688- 3690, 2003.