



Zelig: Statistics for everyone

Authors

Teresa Sancho Vinuesa
tsancho@campus.uoc.es

David Cabanillas Barbacil
dcabanillas@campus.uoc.es

Summary

Zelig is an environment designed for teaching and learning Statistics - a common subject on many academic courses - over the Internet. The contents, appropriately marked-up, are presented to the students according to their various features: the study plan, prior learning and their capabilities

Summary

1. Introduction
2. The Zelig learning environment
3. The basic pillars of the learning environment
 - 3.1 The basics of knowledge
 - 3.2 Technology
4. Conclusions and prospects
5. Sources of information

1. Introduction

La Universitat Oberta de Catalunya -UOC [The Open University of Catalunya]- was born at the end of the 1990's to offer open university study facilities in Catalan. It began teaching the second cycle of Psychopaedogogy and a Diploma course in Business Studies to later increase its offer to other university courses: Humanities, Catalan philology, Documentation, Law, Psychology, Business Administration and Management and I.T. Engineering. It now includes ongoing learning courses, postgraduate Masters' degrees and University Entrance Courses for over-25's. It has recently opened its doors to the Spanish-speaking community through the Planeta-UOC initiative and has started up a Doctorate Programme covering the information society and knowledge, a multidisciplinary, internationally-focussed course.

If, within this context, we were to identify each and every one of the existing subjects - with their corresponding indices - we might be able to profile larger themed areas and at the same time discover transversal subjects of interest to all of them. One of these is of especial interest to us: **Statistics**.

In almost all areas of learning, from engineering and sociology, through to biology and economy, certain issues arise which are hard to describe using mathematical models which are often too complex or, conversely, too simple. In practice, results are often obtained from experience and from a series of well-organised and appropriately-treated data. Statistics studies how to obtain relevant information from a series of data and can become a fundamental tool in the decision-making process. That is why it is a subject studied in many university courses, but always under the same name: "Statistics". In some cases the name is tagged with the context within which it is used ("Statistics applied to human and social sciences") and in others it appears under a different name ("Data Analysis"). Evidently, however, the basic concepts and most frequently-used techniques are common to all of them and differences are to be found fundamentally in the degree of formalisation and in-depth study of certain content and in some contextualisations. In view of this, in 1998 we discussed the idea of creating a **common, adaptive learning environment** for the various disciplines, which would help the student attain the objectives established for the subject for which he/she has enrolled; what we currently have is a prototype.



2. What we call the Zelig learning environment

We must bear in mind that the relationship between the student and the UOC is established through a Virtual Campus, an Internet-based communications environment. The Virtual Campus allows students and lecturers to interact, without having to occupy the same time or space; it also facilitates communication with the other areas within the university itself (the Secretary's office, library, I.T. helpdesk, etc.) through a PC connected to the Internet. It is within this contact that the learning environment we call Zelig was created. This environment is common to all Statistics courses. The environment must integrate space for communications between students and lecturers, content and the scheme of work established for each student. This is why the environment should not adapt only to the profile of the student - prior knowledge and subject he/she wishes to pass - but should adjust depending on his/her behaviour; that is, depending on how he solves the exercises set, the frequency and type of help he requests or the appreciation the lecturer has of his/her progress.



Zelig is the name of a film acted and directed by Woody Allen in 1983. Using a false documentary, Allen introduces us to Leonard Zelig, an imaginary chameleon-like character who can change into anyone he comes into contact with

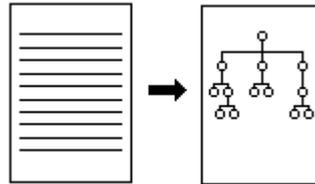
In this sense, once the user has been identified, the system must be able to recognise him/her and allow him/her to initiate the **working session**. Evidently, he/she must be provided with all general information on the subject he/she is studying during the first session (objectives, index, methodology, evaluations) and the scheme of work which must allow him/her to have attained the established objectives at the end of the semester.

The word 'zelig' is taken from the German word 'selig'. Seligkeit, which in English means beatitude or blessedness, is almost synonymous with a state of supreme happiness.

By following a 'Learning by doing' methodology, the student is offered the chance to perform a statistics-related task, the interest of which lies both in the conclusions he/she can draw from it and the processes and tools used. This task is made up of several parts: collection and organisation of data, descriptive analysis of the data, interpretation and meaning in relation to populations and conclusions and decision-making. Each



student's learning process is built up adaptively, based on how he/she resolves these practical tasks, requiring the student to apply the various concepts he/she has assimilated. The teaching strategy for the student to assimilate the necessary content is provided by the scheme of work in which targets, content and the activities necessary to attain these targets are determined in a personalised fashion.



New document structuring

It is, then, important to highlight the fact that this learning process takes place on a network where the different agents involved can interact: students, lecturer and content. Interaction takes place through the resolution of exercises and experimentation in simulation areas, and determines the rhythm of the student's progress.

3. The basic pillars of the learning environment

In terms of the common, adaptive learning environment for different university degrees which include "Statistics" within their syllabus, we have explored different didactic materials for the teaching of quantitative subjects. By paying particular attention to the methodology, the content and the technology, the basic pillars which support the following have been established: the basics of knowledge, the interface, specialised strategies and the necessary technology to make everything work correctly.

3.1 The basics of knowledge

Evidently it is not enough merely to have all the content at one's fingertips to satisfy the various requirements: the content must be designed and stored appropriately for it to be managed. This is why the process of creating the material requires a different construction to conventional subjects within remote teaching and through a virtual learning environment. In particular, given the fact that material provided by authors from very diverse disciplines must co-exist, we must have clearly marked rules which will allow the system to treat them all in a uniform fashion. On the one hand the system must be able to manage all the content and on the other, it must be possible for this content to be shared between the various environments.

the first stage - generating the material - is dealt with by the authors. They have to fill in certain templates created in Microsoft WORD, which will allow them to do so in a structured fashion. These templates clearly indicate how the fields must be filled in and what information is required for each content. These templates can be used to introduce all manner of content from any UOC study which refers to the subject of Statistics.



New document structuring



The important distinction, in relation to the creation of traditional content, comes when the data is introduced: apart from the data itself, the features of the data must also be keyed in (degree of difficulty, expected time for an answer, and other details); this complementary information is used by the system to find the most appropriate datum in each case.

To begin with, three groups of authorship have been established according to the various study plans containing a subject-matter relating to Statistics: Technical Engineering and I.T., Business Studies and Documentation. Two types of core information are used to organise the information: content (conceptual and procedural) and exercises. To create the learning paths we must establish the logical relationships between them, bearing in mind their marks and attributes, which must be designed for them to be appropriately viewed.

3.2 Technology

We shall now show the technology used to create the prototype and we shall introduce the concepts of these new technologies to the reader since in the future they will probably end up becoming part our daily vocabulary, just like technological terms such as Internet, or HTML page.

It is evident that the implementation of the Zelig environment forces us to re-focus both the design and process for creating the teaching materials associated to a study course within a virtual teaching and learning system. In fact, it goes a little further than that : it has meant re-designing the whole subject, both as regards content and the space for communication between lecturers and students, and the level of follow-up and presentation of the content to the students.

We have three main tools available to us to achieve this objective, and this prototype is based on them: the mark-up language, databases and Javascript. It is likely that in the near future all Web applications will combine these elements. Even though each one of them can provide a good range of functions, together they give us all the functions required by any kind of Web application

"Mark-up language"

The mark-up language used by the prototype - XML (eXtensible Markup Language) - provides a structure for the data and allows us to introduce properties, so as to give information about the information, the so-called meta-information. If we can structure the data it is easier to process them.

By creating documents in this new format we can separate the content (XML) from the representation (XSL).

"Data Repository"

Should we need to have access to a large number of stable data, we must use a database. In this case here, we must manage all the materials associated with all Statistics-related subjects, and so we have to use a database.

"Logic"

The prototype's logic has been created using Javascript; this allows us to monitor the students and offer them, on a case-by-case basis, the most appropriate data for their requirements. We are attempting to create a ductile and ergonomic system, which will adapt to the specific requirements of each student.

4. Conclusions and prospects

The work we have done has given us a new perspective on both the creation process and the design of teaching materials required for a virtual teaching and learning system. This perspective is based on the atomisation of contents and on the creation of all the possible links between them. In this way, we can achieve



an adaptive environment which will help personalise the teaching provided and ensure that the course objectives are met.

Technologically we must emphasise the fact that we are working with standard mark-up languages like XML and XSL. This provides us with an internal cohesion of the structure of the content taught in each UOC subject, and an increase in scope in the expectations for the future exchange of knowledge on a global scale, most particularly within the university environment.

The first stage of Zelig focuses on the environment of Statistics. The fact that this structure can be used to cover other areas of knowledge should mean that all information associated with a certain subject should be structured beyond the classical concept based on grouping contents together into subjects. Following this line of research, we might reach a situation where, for example, the mathematical content of a certain subject is studied when necessary for the student to learn a certain concept and not within a specific subject.

The construction and implementation of the prototype has brought certain aspects to the fore, which should be the subject of a more in-depth study: authorship tools to generate and maintain appropriately marked up content, interface design, tools for the lecturer to follow and intervene in the student's learning process and tools for the automatic evaluation of the student.

5. Sources of information

G. Smith and P. College. 1998. Learning Statistics By Doing Statistics. *Journal of Statistics Education* v.6,n.3. <http://www.stat.unipg.it/ncsu/info/jse/v6n3/smith.html>

T. Sancho, T. Meseguer, F. Vallverdú, E. Batlle. 2000. Zelig: Statistics for all. *M/SET International Conference on Mathematics/Science Education & Technology*. [S_Diego_zelig.doc](#)

Maruyama, Tamura, Uramoto. 1999. Java with XML: Creación de sitios web con XML y Java.(Creation of web sites with XML and Java). *Prentice-Hall*. ISBN: 84-8322-242-6

E. Rusty Harold. 1999.XML Bible. *IDG Books*. ISBN: 0-7645-3236-7

Central W3C address (World Wide Web Consortium)

<http://www.w3.org/>

The Sun site on Javascript. <http://java.sun.com/>

IMS (Instructional Management Systems) site. <http://imsproject.org/>

Recommended citation:

SANCHO VINUESA, Teresa; CABANILLAS BARBACIL, David (2001). "Zelig: Statistics for everyone". *Digithum*, issue 3 [article online]. DOI: <http://dx.doi.org/10.7238/d.v0i3.594>