Authors' Information

Jesús Cardeñosa – Department of Artificial Intelligence; Universidad Politécnica de Madrid; Madrid 28660, Spain; e-mail: carde@opera.dia.fi.upm.es

Carolina Gallardo – Department of Artificial Intelligence; Universidad Politécnica de Madrid; Madrid 28660, Spain; e-mail: carolina@opera.dia.fi.upm.es

Álvaro Martín – Department of Artificial Intelligence; Universidad Politécnica de Madrid; Madrid 28660, Spain; e-mail: martin@opera.dia.fi.upm.es

EXPERIENCES ON APPLICATION GENERATORS

Hector Garcia, Carlos del Cuvillo, Diego Perez, Borja Lazaro

***Abstract**: The National Institute for Statistics is the organism responsible for acquiring economical data for governmental statistics purposes. Lisbon agreements establish a framework in which this acquisition process shall be available through Internet, so each survey should be considered as a little software project to be developed and maintained. Considering the great amount of different surveys and all changes produced per year on each make impossible this task. An application generator has been developed to automate this task, taking as a start point the Word or PDF template of a survey, and going through a graphical form designer as all human effort, all HTML, Java classes and Oracle database resources are generated and sent from backoffice to frontoffice servers, reducing the team to carry out the whole set of electronic surveys to two people from non I.T. staff.*

***Keywords**: Software automation, application generators, CASE.*

***ACM Classification keywords**: D2.2. Design Tools and Techniques*

Introduction

Complaining Lisbon agreements concerning e-Government, the Spanish National Institute for Statistics (INE) tackles the problem of translating all economical surveys from paper format into web applications. There exist hundreds of different forms, and for a particular survey, more than one version depending on the kind of target organization, so the required effort to create all infrastructure exceeds not only the capacity of I.T. Department, but the budget to carry out the gigantic task. A previous successful experience on metadata processing from INE and the pilot projects on Java application generation from Technical University of Madrid seem a proper combination to afford the trouble.

The idea consists of taking as a start point the current survey forms in Microsoft Word or PDF format, translating these into a tag based format appropriate for both browser representation and automated processing. This creates some kind of template used as a background for the application. Then a user may define the web form over the background painting components using a designer, and establishes properties for the components from those pre-defined in the designer. Finally only translating these definitions into source code is still to be done.

The technology of generated code shall meet the following requirements:

- HTML 4.01, later substituted by XHTML 1.1 by the research team at UPM, for the web user interfaces.
- XForms 1.0, for the definition of validation rules, with the premise to deploy complete surveys in XForms for future use.
- Java servlets, based on action struts architecture and their corresponding beans.
- Hibernate 3 as database connection tier.
- PDF format as receipt of the answered surveys.

The Architecture

The main goal of the project, beyond any other, was to decrease sensitively the staff, effort and time to market for each survey application, and so, of the whole set of applications. The lack of I.T. professionals in the department in charge of the project also conditions the profile of the target user of the generator.

Four modules were found to be the core of the survey generator:

Format translation tool

As long as the forms corresponding to the different surveys are being created in other departments, the format and composition developed for hard copies is not valid for automated processing, some tool to extract the contents from Word and PDF files and export them into tagged files, closer to web requirements and much more appropriated for processing.

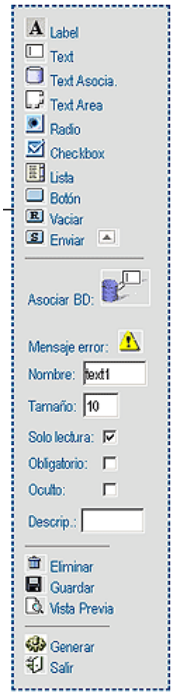
For this purpose several options were evaluated. At first the best choice seemed to develop a specific translator, in Java language. The wide support for PDF processing available supposed a great advantage, but by that time the number of API or information about accessing Word files, especially about the structure of these files, was very poor. Only some arising APIs, such as Apache's POI, were available, so finally the decision went on a third party product, and then develop only the integration to the system.

Survey definition tool: the editor

Once the source document has been translated in a processable format, and before proceeding to its publishing, the system allows users to define the forms for the survey. Of course, there is not enough information in the templates, but visual aspect. In this sense, the captured survey is shown to the user as background in a screen, in which he may add or remove components that will later compose a web form.

For each field in the survey the editor allows to define some specific features, such as data type, length, etc. In case of combo boxes or lists, it is possible to define the valid values list. Also constraints have been implemented, such as date formats, decimal and thousands separator, ranges, allowed data sets, etc.

The components available while designing forms are the following:



Nombre: text1

Tamaño: 10

Solo lectura: ☑

Obligatorio: □

Oculto: □

Descrip.:

Figure1.- Toolbar

Label: consists of a read only text. The user may define component name, text, font type, size and color, bold and italics style and background color.

Text: it is a read/write field containing characters, and it is possible to configure component name, length, maximum capacity, data type (string, date, time, year, day, month, float, double, integer, positive, negative, long) for validation rules to be applied. If the validation finds that the content of the field does not match the data type, the user shall be advised, so the application allows the user to set an error message to be displayed in a dialog. It is possible to apply some modifiers to the text fields, such as mandatory, read only, hidden and calculate value automatically. In case of selecting calculated field, a calculator is shown to define the formula. A formula may contain both values or fields in the same survey. There is also a description of the field to be shown as a hint. For some specific surveys, with repetitive contents, such as tables with a row per city or state, it is necessary to associate the field with a column in the table where data is to be stored.

Text from database: it is a read/write field associated to a column in a table from a database. The purpose is to set a default value for the field when the form is loaded. The field is processed as a text field once the form is submitted. It is very useful when defining, for instance, headers in a form, with the name, address and essential data for the final user. Possible configurations are identical to text fields.

Text area: consists of a text field with several rows. The user may define the size in terms of rows and columns, and data type, read only and mandatory options are available. For this component the validation rules may be disabled.

Radio, check boxes and lists: a new feature is added regarding the previous components: an interface to manage the choices and value for each choice in the available options, and establish a default selection if any.

Buttons: it is possible to draw buttons and associate a number of actions to them. Currently a button can save the form, add or delete items to repetitive contents (such as tables), validate the form and generate the errors list, or open the help page. Also specific separate components are available to create reset and submit buttons.

Of course a survey can be defined in several working sessions, so the functionalities for saving a survey in the current definition state, open a saved survey, and delete are available, as well as preview of the current survey. Finally, in the bottom of the toolbar, generate survey and exit commands are located.

If the user has already defined other surveys, the system enables the import functionality, so that the components from other surveys can be brought. This feature is especially useful to create surveys that are simple modifications from the previous year format, or to create the same survey in a different language, so only some modifications on descriptions and hints are needed.

Most surveys are so long that page breaks are needed. The editor allows defining these line breaks, which also are considered when generating the PDF receipts.

Application generator

The application generator retrieves the form already defined with the previous tool, starting a analysis process, in which the definition data are separated into modules, and afterwards the different application components are generated as a set of XHTML pages that are a composition starting with the original HTML from the PDF or Word, adding the information defined in the editor, and some GEN specific attributes for some tags that allow further processing. These attributes are used mainly to show the specific user information (i.e. tags with organization information, or the form with values that have been saved in a previous session).

Then the XForm files needed to validate the data in the forms are built to be used in web and bulk load processes, plus the resources for XHTML, such as background images, help files, etc.

All these infrastructure is stored in the database, and when a request to fill in a form is produced a set of servlets, action struts and filters that parse the XHTML page to be showed, together with the required resources, get the result.

To generate PDF receipts from the submitted forms XSL:fo is used. As long as the forms are, after the process, HTML pages, a previous transform has to be done. Fo processes basically XML documents, and HTML does not provide the closing tags, for instance <BR>, so the document is analyzed and translated into XHTML before processing. The PDF shall have same number of pages as defined in the survey using the page breaks.

Publishing tool

The first approach to publish all the resources generated by GEN was to send to the production server all classes, JSPs, servlets, images, etc. It is a remote server in a different network, and should be configured for allowing the access to file structures, execute compilations, modify Tomcat settings, … at the same time that the network devices where also reconfigured. Of course this is not an affordable task, obviously a software tool of this kind cannot cause all this changes.

Finally, this led us to refactor all design, and instead generating all the mentioned resources, two meta servlets where produced, so that one of them is capable to draw any form and the other one is capable to receive (after submitting) any kind of form. This reduces sensibly the publishing process, now only access to the database is required, and all resources are stored there. The database is accessed through the corporate intranet.

Case Study

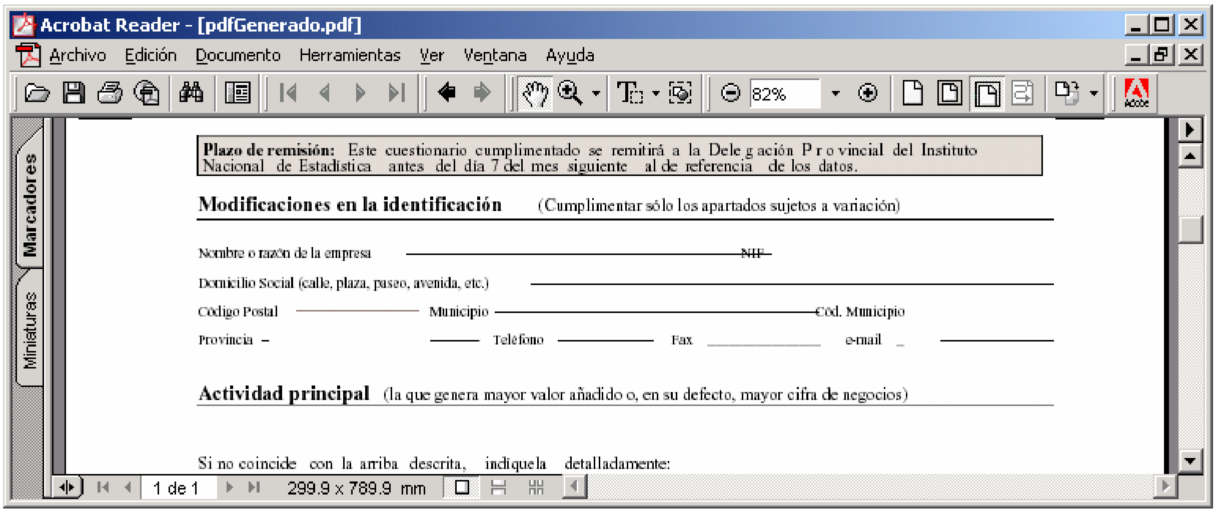
Probably the best way to describe how GEN works and the simplicity of the process to generate the new application to manage a survey is to follow an example.

First step: translating a PDF survey into HTML

The first of all, we need the file containing the survey designed by the corresponding department. Usually surveys have been translated into PDF, anyway this is an immediate process, usually covered by Adobe Acrobat Writer or Distiller, as well as by any tool complaining the standard.

This translation is based on an external tool, so we do not mind how the translation is done, the only important issue is the quality of the HTML we get after the translation. GEN only needs as input a HTML complaining the W3C standard. We always avoid HTML resulting from Microsoft Word exports, FrontPage, etc.

In this case study, we are using a translation tool which converts PDF into an HTML composed of layers. Each component of the document is set into a different layer, and the absolute position for each layer is written in the code. This is very useful for future processing of the template, while painting the components in the proper location over their relative texts and tables.


Acrobat Reader - [pdfGenerado.pdf]
Archivo Edición Documento Herramientas Ver Ventana Ayuda
82%
Marcadores
Miniaturas
1 de 1
299.9 x 789.9 mm

Plazo de remisión: Este cuestionario cumplimentado se remitirá a la Delegación Provincial del Instituto Nacional de Estadística antes del día 7 del mes siguiente al de referencia de los datos.

Modificaciones en la identificación (Cumplimentar sólo los apartados sujetos a variación)

Nombre o razón de la empresa ______ NIF

Domicilio Social (calle, plaza, paseo, avenida, etc.) ______

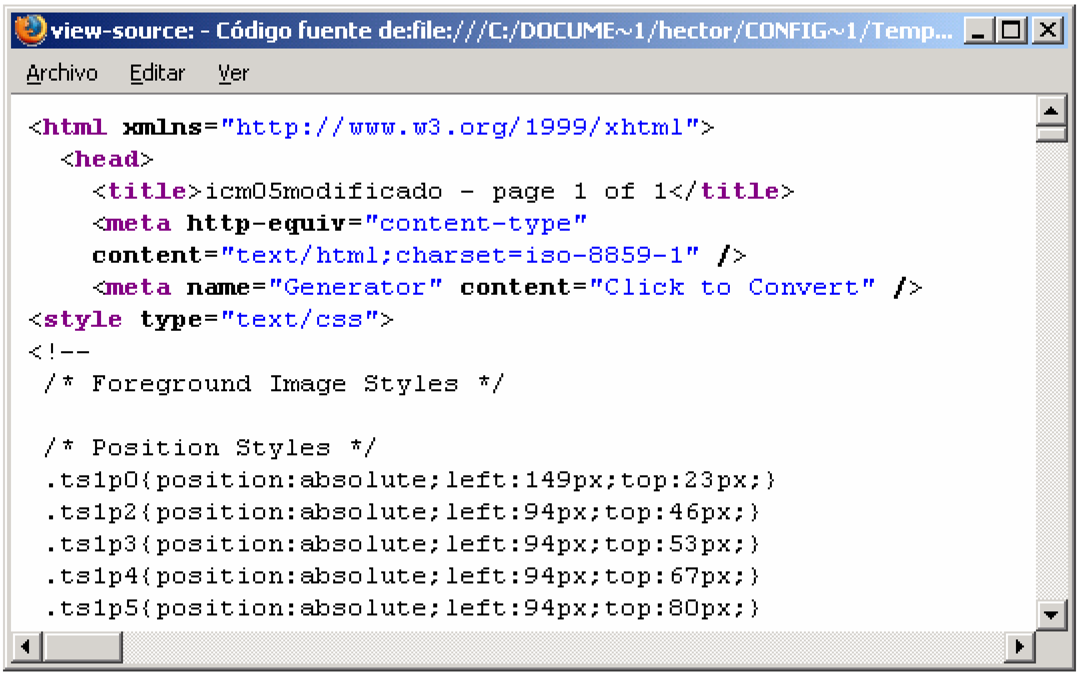
Código Postal ______ Municipio ______ Cód. Municipio

Provincia ______ Teléfono ______ Fax ______ e-mail ______

Actividad principal (la que genera mayor valor añadido o, en su defecto, mayor cifra de negocios)

Si no coincide con la arriba descrita, indíquela detalladamente:

Figure 2.- The original PDF survey


view-source: - Código fuente de:file:///C:/DOCUME~1/hector/CONFIG~1/Temp...
Archivo Editar Ver

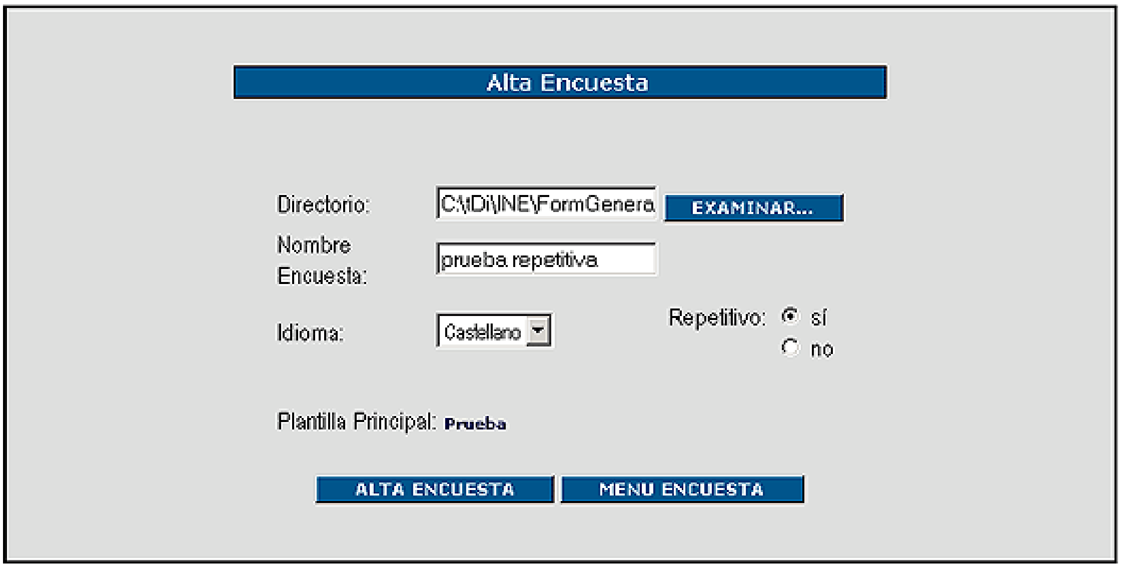
```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>icm05modificado - page 1 of 1</title>
    <meta http-equiv="content-type"
    content="text/html;charset=iso-8859-1" />
    <meta name="Generator" content="Click to Convert" />
<style type="text/css">
<!--
 /* Foreground Image Styles */

 /* Position Styles */
 .ts1p0{position:absolute;left:149px;top:23px;}
 .ts1p2{position:absolute;left:94px;top:46px;}
 .ts1p3{position:absolute;left:94px;top:53px;}
 .ts1p4{position:absolute;left:94px;top:67px;}
 .ts1p5{position:absolute;left:94px;top:80px;}
```

Figure 3.- The HTML obtained from the PDF showing the layer positions

Second step: design the form for a survey

Now GEN user shall define a form to capture the data corresponding to the survey. The HTML from the previous step is considered as a simple background to help the user drawing all components. This is the main task while using the editor, but also some other functions are to be completed: buttons to send, save or step through the forms, page breaks for viewing and generating PDF receipts, or dependencies and constraints specification.



Alta Encuesta

Directorio: C:\tDi\INE\FormGenera

Nombre Encuesta: prueba repetitiva

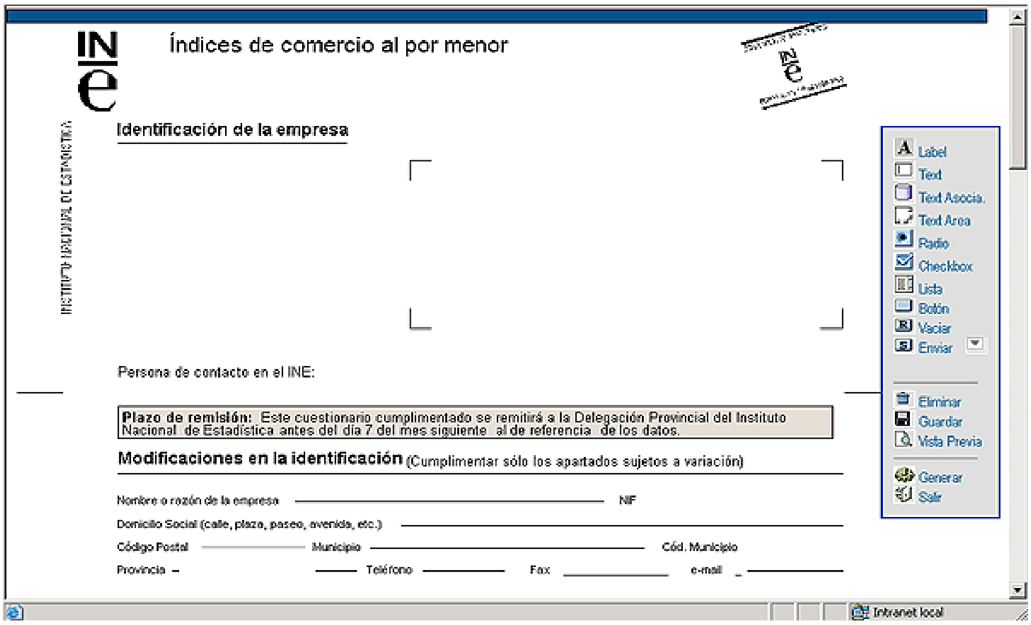
Idioma: Castellano

Repetitivo: ● sí ○ no

Plantilla Principal: Prueba

Figure 4.- Creating a new survey

First of all the typical screen where user is going to define the working directory, survey name and type. GEN allows to define multilingual surveys, so that the language specification is also important for further publishing. Each survey may be also converted in a template that may be used to create compositions in which a survey is composed of modules that are really survey templates. This is very useful to define headers or footers, or to include forms common to many applications.



Índices de comercio al por menor

Identificación de la empresa

Persona de contacto en el INE:

Plazo de remisión: Este cuestionario cumplimentado se remitirá a la Delegación Provincial del Instituto Nacional de Estadística antes del día 7 del mes siguiente al de referencia de los datos.

Modificaciones en la identificación (Cumplimentar sólo los apartados sujetos a variación)

Nombre o razón de la empresa ____________ NIF

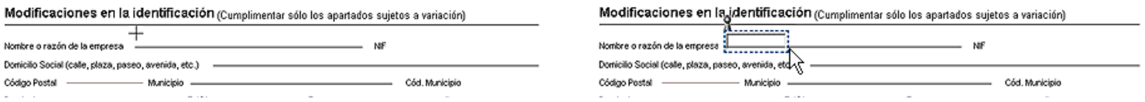
Domicilio Social (calle, plaza, paseo, avenida, etc.) ____________

Código Postal ________ Municipio ____________ Cód. Municipio

Provincia – ________ Teléfono ________ Fax ________ e-mail _ ________

Figure 5.- The survey is loaded in the editor to create the form

Once the survey is loaded in the editor, the user may define the components and its location over the background. The layer information from the HTML obtained in the first step is very important now; when a user draws a component in the form, the editor calculates the position trying to associate the component to a field in the survey, matching them depending on the proximity.



Modificaciones en la identificación (Cumplimentar sólo los apartados sujetos a variación)

Nombre o razón de la empresa ______ NIF

Domicilio Social (calle, plaza, paseo, avenida, etc.) ______

Código Postal ______ Municipio ______ Cód. Municipio

Modificaciones en la identificación (Cumplimentar sólo los apartados sujetos a variación)

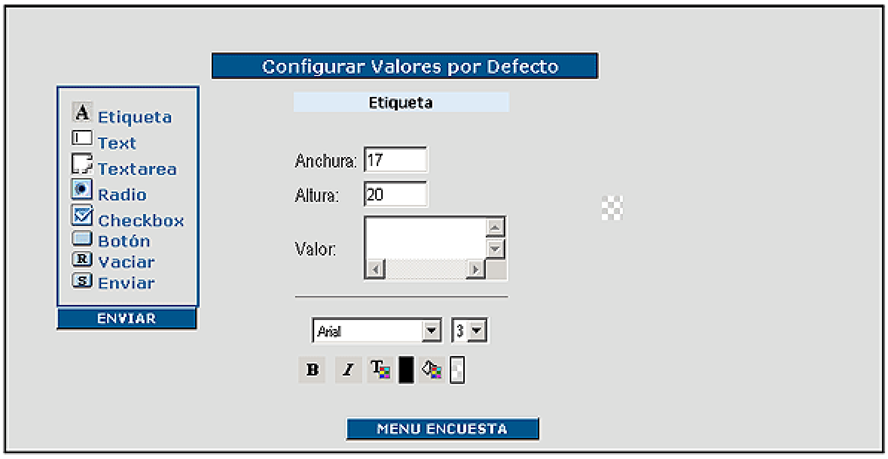
Nombre o razón de la empresa ______ NIF

Domicilio Social (calle, plaza, paseo, avenida, etc.) ______

Código Postal ______ Municipio ______ Cód. Municipio

Figure 6.- Drawing a component from the toolbar

Components are managed as usual in an IDE, including the possibility of copy, paste, align vertically or horizontally, size, resize or set equal size for several of them, defining default values for each component type, etc.



Configurar Valores por Defecto

Etiqueta

Anchura: 17

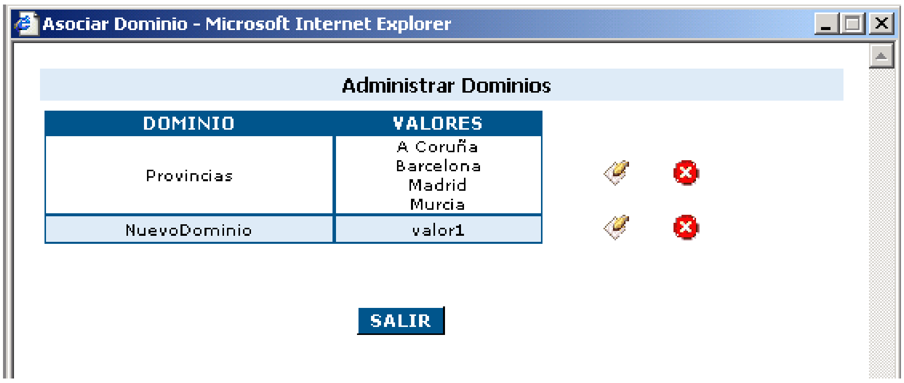
Altura: 20

Valor:

Arial 3

Figure 7.- Defining default values for label components in a survey

In some cases it is very interesting to define a domain of values for some components, specially in case of combo boxes or lists, in the figure below a example on how to define a fixed domain for states is shown. The possibility of recovering dynamically the domain from a table in a database has been also implemented.



Asociar Dominio - Microsoft Internet Explorer

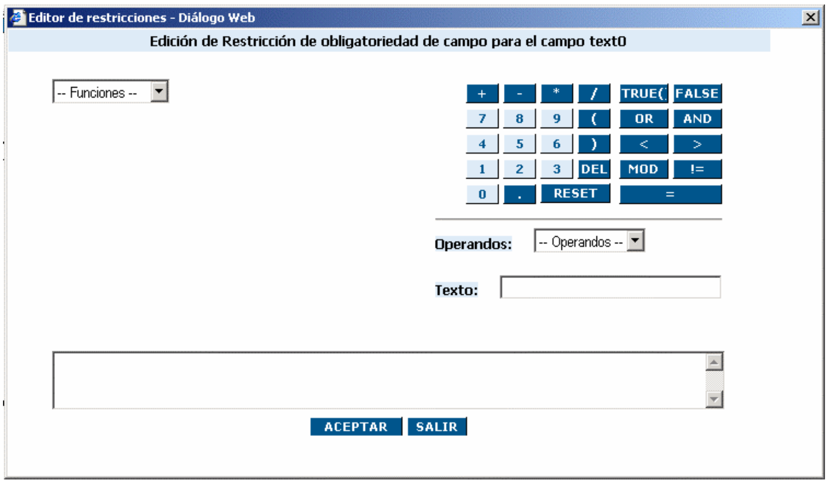
Administrar Dominios

DOMINIO	VALORES
Provincias	A Coruña Barcelona Madrid Murcia
NuevoDominio	valor1

Figure 8.- Domain definition window

Another interesting functionality while defining surveys is to set some kind of constraint to the values of the different fields. The user may define both individual or group constraints. Individual constraints are to be applied to the value of a field, while group constraints are applied to the set of values of some fields. In any case, the constraints may obly to fill the field (*mandatory*) or define restrictions on the value of the field if any, through formula specification, as shown in the figure below. In the combo box the user may select some common predefined functions.

After some time defining the form associated to a survey, the user may have added and removed several components (in fact, usually, dozens of them). Each new component is named automatically by the editor and introduced in the tab order. The user may alter both, the given name to make the application comprehensive (very recommendable) and also the tab order, so that the final user may navigate through the fields with the tab control, which is one of the most accepted (and requested) features from final users.



Editor de restricciones - Diálogo Web

Edición de Restricción de obligatoriedad de campo para el campo text0

Operandos: -- Operandos --

Texto:

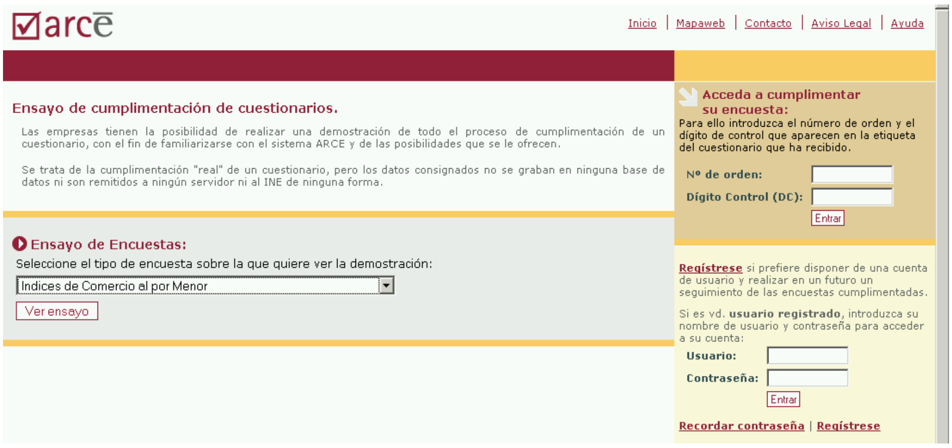
Figure 9.- Constraint definition

Not only the form to publish the survey is important, but also to provide the users an identical aspect and information to reduce the difference to the hard copy surveys. In the same terms the background is the same that in that case, the designer provides the chance to add help buttons to the form, then these buttons may be linked to a help file in HTML format that usually contains the same help provided to non Internet users.

The results: an external user fulfilling a survey

When the process is finally finished, the user gets a result which may be considered in two perspectives:

- An application that supports the publication and collection of data from the web, on a surprisingly low cost and effort, and which is integrated in the portal of the organization.
- A service for the information providers, those are obliged by law to compliment the surveys, and may now carry out this task easily.



Ensayo de cumplimentación de cuestionarios.

Las empresas tienen la posibilidad de realizar una demostración de todo el proceso de cumplimentación de un cuestionario, con el fin de familiarizarse con el sistema ARCE y de las posibilidades que se le ofrecen.

Se trata de la cumplimentación "real" de un cuestionario, pero los datos consignados no se graban en ninguna base de datos ni son remitidos a ningún servidor ni al INE de ninguna forma.

Ensayo de Encuestas:

Seleccione el tipo de encuesta sobre la que quiere ver la demostración:

Indices de Comercio al por Menor

Acceda a cumplimentar su encuesta:

Para ello introduzca el número de orden y el dígito de control que aparecen en la etiqueta del cuestionario que ha recibido.

Nº de orden:

Dígito Control (DC):

Regístrese si prefiere disponer de una cuenta de usuario y realizar en un futuro un seguimiento de las encuestas cumplimentadas.

Si es vd. **usuario registrado**, introduzca su nombre de usuario y contraseña para acceder a su cuenta:

Usuario:

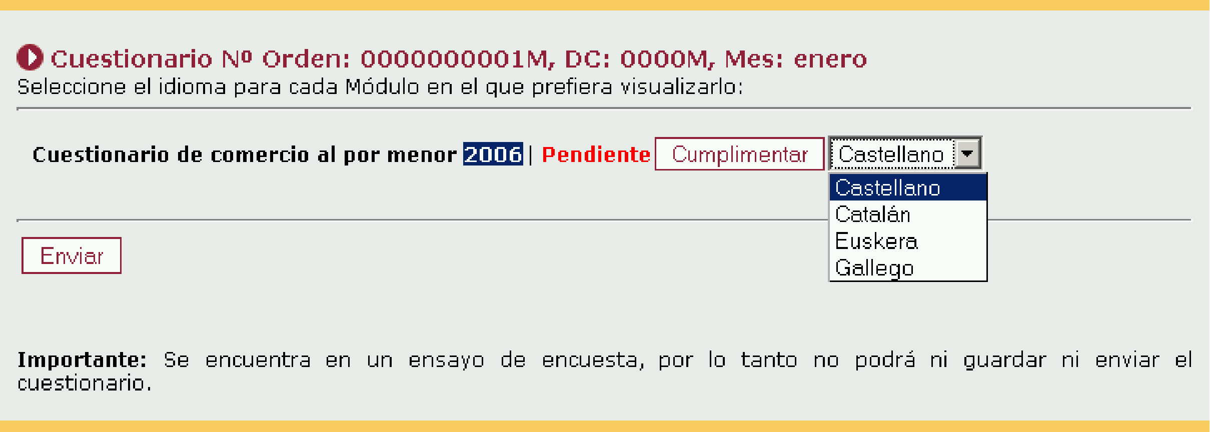
Contraseña:

arcē
Inicio | Mapaweb | Contacto | Aviso Legal | Ayuda
Ver ensayo
Entrar
Entrar
Recordar contraseña | Regístrese

Figure 10.- The web site allows not only filling real surveys, but making tests

The applications generated by GEN are integrated in the web site through shared tables in an Oracle database, and the development of a section in which the user may select the survey in which he is interested. The access to the section is based on the typical user and password, that has been replaced by an order number and control digit. This information is sent to each organization together with customized instructions and paper surveys via mail as official notifications, reusing the existing infrastructure and the old procedure, which is mandatory to maintain.

When a survey is generated, the user may define whether the survey shall exist also in test mode. This allows final users to practice with real surveys without writing results or submitting information. The only difference between tests and real surveys is a parameter switching between them.



Cuestionario Nº Orden: 0000000001M, DC: 0000M, Mes: enero

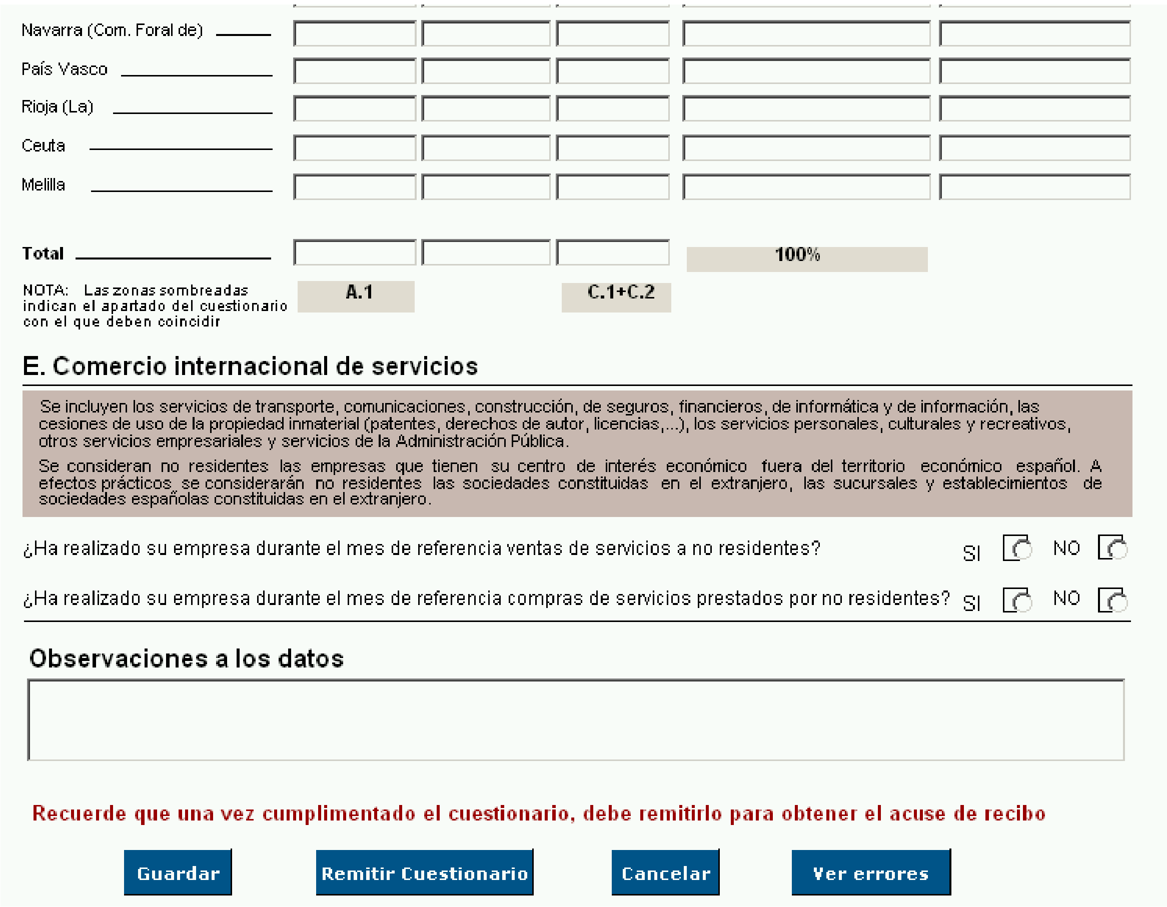
Seleccione el idioma para cada Módulo en el que prefiera visualizarlo:

Cuestionario de comercio al por menor 2006 | **Pendiente** Cumplimentar Castellano

Importante: Se encuentra en un ensayo de encuesta, por lo tanto no podrá ni guardar ni enviar el cuestionario.

Figure 11.- Language selection. GEN generates multilingual applications

The surveys can be displayed in all Spanish official languages as mandatory by law, but the data are considered to correspond to the same survey, this is logical until a problem arises: the staff in charge of collecting data and extracting statistical information do not use to speak all of them. It is considered a great lack not for the application but for the legal procedure.



Navarra (Com. Foral de)					
País Vasco					
Rioja (La)					
Ceuta					
Melilla					
Total				100%	
NOTA: Las zonas sombreadas indican el apartado del cuestionario con el que deben coincidir	A.1		C.1+C.2		

E. Comercio internacional de servicios

Se incluyen los servicios de transporte, comunicaciones, construcción, de seguros, financieros, de informática y de información, las cesiones de uso de la propiedad inmaterial (patentes, derechos de autor, licencias,...), los servicios personales, culturales y recreativos, otros servicios empresariales y servicios de la Administración Pública.

Se consideran no residentes las empresas que tienen su centro de interés económico fuera del territorio económico español. A efectos prácticos, se considerarán no residentes las sociedades constituidas en el extranjero, las sucursales y establecimientos de sociedades españolas constituidas en el extranjero.

¿Ha realizado su empresa durante el mes de referencia ventas de servicios a no residentes? SI ○ NO ○

¿Ha realizado su empresa durante el mes de referencia compras de servicios prestados por no residentes? SI ○ NO ○

Observaciones a los datos

Recuerde que una vez cumplimentado el cuestionario, debe remitirlo para obtener el acuse de recibo

Figure 12.- A piece of the final presentation of a survey on commerce

In figure 12 a fragment of the web showing a survey is presented. Note the final alignment of the fields, the different component types and the buttons in the bottom, offering the functions of *Save*, *Send form*, *Cancel* and *Show errors*.

Before sending the final answer to a survey, the user may save any number of times, and before submitting may query the errors in the different fields in order to make corrections. Sending a correct answer is mandatory by law.

The final reply of a user to a survey is stored in XML format into the database, and will be later exported to the analysis application.

This is the common process for most users, however, a number of enterprises, mainly holdings, do need to fill lots of surveys that they may automate by using their information systems. A complementary application has been developed allowing this kind of users replying massively to the surveys through bulk loads. This application, called G2G, receives XML files containing several surveys, validates the data using the XForms from GEN, after formatting incorrect data or truncating data too long for the precision in database.

Conclusion

Sometimes the public organisations shall face political commitments with little help from the Government, same budgets and staff, allowing the development of beautiful projects in which technology makes life easier. Probably this project had lost the chance if there where enough staff to develop each application as usual software projects.

A tool has been developed that allows a non technical user, with no knowledge on software architectures, design, HTML, nor Java, struts or databases, carry out the task of creating applications. Of course this does not replaces technical staff, but get them to develop only technical high level tasks.

The fact is that a complex technology, that has been tested by Technical University of Madrid, and that shuns from the typical theoretical aspects from code generation, trying to open a new spectrum of possibilities, has been applied successfully to a specific application domain demonstrating that the approach is good and feasible.

For public administrations, the project allows a rise in the corporate image, and the observance of the Lisbon agreements in one of the most difficult aspects it was facing to.

The Spanish National Statistics Institute has been capable to meet the goals described in its services white-paper, minimizing the impact in the staff and budget, and reducing sensitively the time to market for each survey, which is available in Internet in less than a week, so that the complete process increases a minimun percentage from the previous one, with no digital survey. Also the time in which all data has been collected is lower, because there is no need to transcript all data to a computer.

Collaborators that fill in the forms, now have the chance to select the preferred choice: digital or paper surveys, knowing that digital ones provide the proper mechanisms to automate the calculations, report possible errors reducing the time to finish the job, etc.

The success in figures: in the first two months more than 10% of the surveys where submitted via Internet, more than 600 per day are being received, and 7 multilingual surveys have been generated; 168.433 organizations are working with the generated applications.

Bibliography

[Alves2002] Alves, L., von Staa, A., 2002. A construction process for artifact generators using a CASE Tool. *Proceedings of Workshop on Generative Programming 2002.* Austin, Texas, USA, pp. 7-10.

[Arisholm1998] Arisholm, E. et al. 1998. Incorporating Rapid User Interface Prototyping in Object-Oriented Analysis and Design with Genova. *The Eighth Nordic Workshop on Programming Environment Research.*

[Batory 1994] Batory, et al. 1994. The GenVoca Model of Software-System Generators. *IEEE Software*, 0704-7459/94.

[Boehm2000] Jongmoon, B., Boehm, B., 2000. Empirical Analysis of CASE Tool Effects on Software Development Effort. *Technical Report University of South Carolina. Center for Software Engineering.*

[Cuvillo2004a] Cuvillo, C. et al. 2004. Generation tool for DBMS focused applications. *Applied Computing 2004*. Lisbon. Portugal.

[Cuvillo2004b] Cuvillo, C. et al. 2004. Multiplatform web applications generated from relational data models. (In spanish). *CISCI 2004*. Florida, USA.

[Calejo2002] Calejo, M. et al. 2002. Web Application Maker - A model based approach to web database development. *6th International Conference on Enterprise Information Systems.*

[Grønbæk 1991] Grønbæk, K.et al. 1991. ApplBuilder - an Object-Oriented Application Generator Supporting Rapid Prototyping. *The 4th International Conference on Software Engineering & Its Applications*.

[Schmidt2000] Schmidt, D., et al., 2000. POSA2: Pattern -Oriented Software Architecture: Patterns for Concurrent and Networked Objects.

Authors' Information

Hector Garcia - Contrated Professor. Technical University of Madrid. E.U. Informática. Ctra. de Valencia Km. 7. E28031 Madrid. e-mail: hgarcia@eui.upm.es

Carlos del Cuvillo - Associate Professor. Technical University of Madrid. E.U. Informática. Ctra. de Valencia Km. 7. E28031 Madrid. e-mail: ccuvillo@eui.upm.es

Diego Perez -. Consultant. Technical University of Madrid. E.U. Informática. Ctra. de Valencia Km. 7. E28031 Madrid. e-mail: dperez@tdi.eui.upm.es

Borja Lazaro - Technician, Group leader. Technical University of Madrid. E.U. Informática. Ctra. de Valencia Km. 7. E28031 Madrid. e-mail: blazaro@eui.upm.es

MODELING AND ANNOTATING THE EXPRESSIVE SEMANTICS OF DANCE VIDEOS

Balakrishnan Ramadoss, Kannan Rajkumar

Abstract: Dance videos are interesting and semantics-intensive. At the same time, they are the complex type of videos compared to all other types such as sports, news and movie videos. In fact, dance video is the one which is less explored by the researchers across the globe. Dance videos exhibit rich semantics such as macro features and micro features and can be classified into several types. Hence, the conceptual modeling of the expressive semantics of the dance videos is very crucial and complex. This paper presents a generic Dance Video Semantics Model (DVSM) in order to represent the semantics of the dance videos at different granularity levels, identified by the components of the accompanying song. This model incorporates both syntactic and semantic features of the videos and introduces a new entity type called, Agent, to specify the micro features of the dance videos. The instantiations of the model are expressed as graphs. The model is implemented as a tool using J2SE and JMF to annotate the macro and micro features of the dance videos. Finally examples and evaluation results are provided to depict the effectiveness of the proposed dance video model.

ACM Classification: J.5 [Arts and Humanities]: Performing Arts; H.5.4 [Information Interface and Presentation]: Hypertext/Hypermedia.

Keywords: Agents, Dance videos, Macro features, Micro features, Video annotation, Video semantics.

1. Introduction

Dance data is essentially multimedia by nature consisting of visual, audio and textual materials. Dance video modeling and mining depends significantly on our ability to recognize the relevant information in each of these data streams. One of the most challenging problems here is the modeling of the dance video semantics such that the relevant semantics are consistent with the perception of the real world.

The classical and folk dances are the real cultural wealth of a nation. In India, the most important classical dances are *Bharathanatyam, Kadak, Kadakali, Kuchipudi* and *Manipuri* (Saraswathi, 1994). Traditionally, dance learners perform dance steps by observing the natural language verbal descriptions and by emulating the steps of the choreographers. Therefore, the properly annotated dance videos will help the present and future generations to learn dance themselves and minimize the physical presence of the choreographers.