

Enunciat de TFC Bases de dades relacionals

Disseny i implementació de la base de dades d'un: sistema de votació ciutadana a nivell Europeu a través d'Internet

Descripció del projecte

L'objectiu principal d'aquest treball és consolidar els coneixements adquirits al llarg dels estudis pel que fa a assignatures, entre d'altres, de Base de Dades (BD). Amb aquest objectiu, el treball proposat consisteix en implementar un sistema de base de dades per a donar resposta a una petició que ens ha fet la Comunitat Europea.

La Comunitat Europea, dins de la partida pressupostària destinada a fomentar la participació ciutadana dins de l'àmbit polític Europeu, ha decidit obrir un concurs públic per a rebre propostes sobre el disseny d'una BD, que els hi serveixi de magatzem d'informació per a la futura aplicació de votacions ciutadanes a través d'Internet que volen implementar.

El marc de la col·laboració amb la UOC se centrarà només en el disseny de la BD, ja que l'aplicació de gestió es desenvoluparà en una segona fase del pla de sistemes d'informació de la Comunitat Europea.

A nivell general, aquesta base de dades ha de guardar tota la informació necessària per a dur a terme votacions ciutadanes a través d'Internet, així com per a guardar l'històric de les mateixes. Tot això inclou entre d'altres: les dades bàsiques de les votacions, els recomptes final dels vots de cada opció, els censos electorals que participen en cada votació, etc.

El sistema a dissenyar ha de permetre emmagatzemar tota la informació comentada anteriorment i permetre generar les consultes més habituals que es realitzaran. Addicionalment a aquest funcionament, la base de dades s'haurà d'encarregar de precalcular i emmagatzemar diversa informació estadística, tal i com es detalla més endavant en els requisits del mòdul estadístic.

Requisits funcionals del sistema

En el plec tècnic del concurs que ha tret la Comunitat Europea es defineixen els diferents requisits funcionals que ha de complir la base de dades d'aquest sistema de votacions ciutadanes. De forma resumida, els llistem a continuació:

[R1] El model ha de permetre guardar totes les dades associades a una votació, aquestes serien com a mínim:

- títol de la votació, per a la seva identificació ràpida per part dels ciutadans
- la descripció de la votació i de les diferents opcions que podran triar els ciutadans (cada votació tindrà un mínim de dues opcions a triar i no hi haurà un màxim establert).
- interval de dates entre les quals estarà oberta la votació dels ciutadans.
- la data de publicació de la votació (serà sempre un mínim d'una setmana abans que s'obri la votació per part dels ciutadans per tal de deixar un període de reflexió).
- Cens o censos electorals habilitats a votar en funció de l'àmbit de la votació. Val a dir que una votació podrà tenir més d'un cens electoral assignat. Per exemple, podríem tenir una votació que afectés a tots els ciutadans de dos països de la Comunitat Europea o de tota la comunitat o d'una o més regions concretes o d'un o més pobles o ciutats.
- El nom del president de la mesa electoral virtual, serà obligatori en tots els casos.
- L'adreça web opcional on hi haurà explicacions més detallades de la votació.

[R2] El model haurà de permetre guardar el nombre de persones amb capacitat de vot en el moment del tancament de la mateixa, els vots totals que s'han emès realment dins de cada votació, així com el percentatge de participació. També caldrà emmagatzemar el nombre de vots que ha tingut cadascuna de les opcions i el percentatge que aquests suposen respecte al total.

[R3] El sistema ha de permetre la gestió dels diferents censos electorals, concretament hi haurà quatre tipus de cens electorals:

- CENS GLOBAL: inclou a tots els ciutadans amb més de 18 anys corresponents a tots els països de la comunitat europea. Només hi haurà una instància d'aquest cens.
- CENS ESTATAL: correspon a tots els ciutadans d'un determinat país de la Comunitat Europea, en aquest cas s'ha de guardar com a metadada el codi del país en qüestió segons la codificació ISO 3166-1 alfa 2 (http://www.iso.org.es/mk.gd/iso/country_codes/iso-3166-1_decoding_table.htm). Hi haurà una instància d'aquest cens per a cada país de la comunitat.
- CENS REGIONAL: inclou a tots els ciutadans d'una regió concreta d'Europa, com poden ser els *landers* alemanys o les autonomies dins d'Espanya. En aquest tipus de cens es guardarà també la dada de a quin país pertany i el nom de la regió en qüestió.
- CENS LOCAL: Aquest tipus de cens inclourà a tots els ciutadans d'un determinat poble o ciutat d'Europa. És molt similar a l'anterior tot i que a més del codi de país i de regió inclourà la metadada del nom del poble o ciutat al que fa referència.

[R4] De la mateixa manera que una votació pot tenir més d'un cens electoral associat, un ciutadà pot pertànyer a més d'un cens també, per exemple, un ciutadà d'Antequera estaria com a mínim a quatre censos:

- al cens europeu de tipus GLOBAL
- al cens de tipus NACIONAL assignat a Espanya
- al cens de tipus REGIONAL assignat a Andalusia
- i a un cens de tipus LOCAL assignat a la ciutat d'Antequera.

Evidentment s'ha de controlar que una persona física, encara que pugui estar en diferents censos, només pugui emetre un únic vot dins d'una votació concreta. Un cop emès aquell vot no es podrà anul·lar ni rectificar.

[R5] Per a mantenir el secret de les votacions la BD només guardarà un codi aleatori associat a cada ciutadà, així com l'edat del mateix. Aquest codi actuarà com a àlies del ciutadà, permetent que aquest emeti vots de forma anònima, és a dir, a la nostra BD no es guardarà ni el nom del ciutadà, ni cap altra dada que pugui violar el dret a la votació secreta,

aquestes dades es guardaran xifrades en un altra base de dades separada físicament i controlada per un departament de seguretat especial i per tant no entrarien dins del disseny que ens demanen. El que sí que guardarà la nostra BD per a cada ciutadà serà: el codi de país, de regió i de localitat, amb finalitats estadístiques i també per a determinar la pertinença del ciutadà a un determinat cens.

[R6] L'aplicació haurà de disposar, com a mínim, de les funcionalitats següents, tot complint amb els requisits expressats prèviament:

- Procediments d'ABM (Alta + Baixa + Modificació) de les votacions i de la seva informació associada: opcions, vots emesos, censos associats, etc...
- Procediments d'ABM dels diferents censos dins de l'àmbit Europeu.
- Procediments d'ABM dels ciutadans i de la seva assignació als censos
- Procediments de consulta que permetin obtenir:
 - a. Donat un país per paràmetre: el llistat de totes les votacions que s'hi han produït (és a dir, que han tingut assignat algun cens pertanyent a aquell país o al cens global). En aquest llistat caldrà retornar la següent informació bàsica:
 - títol de la votació
 - data d'inici i de fi de la votació
 - nombre total de ciutadans dins dels censos associats a la votació
 - estat en que es troba la votació, p.e : en edició, pendent d'obertura, oberta, finalitzada.
 - i si la votació ha estat finalitzada: percentatge de participació total, opció guanyadora i percentatge de vots de l'opció guanyadora. En cas que no estigui finalitzada hauria de sortir un "-" en aquests camps.

Tot això ordenat cronològicament de forma ascendent per la data d'inici de la votació.

- b. Llistat de les 10 votacions on hi ha hagut més diferencia percentual de vots entre l'opció més votada i la menys votada, indicant:

- títol de la votació
- data d'inici i de fi de la votació
- opció més votada i el seu percentatge respecte el total
- opció menys votada i el seu percentatge respecte el total
- diferència percentual entre les dues opcions anteriors

Ordenat descendentment per la diferència percentual de les dues opcions.

- c. El mateix llistat que l'anterior però en aquest cas les 10 votacions que han tingut menys diferència percentual de vots entre l'opció més votada i la menys.
 - d. Donat el codi aleatori que identifica un ciutadà, el llistat de tots els censos als que pertany. Aquesta funcionalitat servirà per a que un ciutadà pugui consultar a quins censos pertany.
 - e. Donat el codi aleatori que identifica un ciutadà, el llistat de les votacions en les que ha participat i en les que no. Aquesta funcionalitat servirà per a que un ciutadà pugui consultar el seu històric de participació en votacions per pantalla.
- Es valorarà la implementació d'altres procediments o funcionalitats que puguin ser d'utilitat, sempre i quan estiguin documentades i consensuades amb el consultor.

[R15] **Mòdul estadístic**, una part molt IMPORTANT del treball és la implementació d'un mòdul estadístic que s'ha d'alimentar a partir dels procediments que implementin les funcionalitats esmentades, per tal d'oferir les dades següents en temps constant 1, és a dir, fer una *SELECT* sobre un registre d'una taula (que no sigui una vista calculada o materialitzada, ni fent servir funcions d'agregats : *sum*, *avg*, etc .. amb *group by*).

El mòdul estadístic haurà de donar resposta a les consultes següents:

1. Donat un país i un any concret: el nombre de votacions finalitzades que s'hi han produït.
2. Donat un cens i un any concret: el valor mitjà del percentatge de participació, tenint en compte totes les votacions associades a aquell cens electoral que han finalitzat durant aquell any.
3. Donat un cens i un any concret: el nombre de votacions que ha tingut associades aquell cens durant aquell any.
4. Donat un any concret: la votació que més participació ha tingut.
5. Donat un any concret: la localitat europea que més vots han emès els seus ciutadans.
6. Donat un any concret: percentatge de ciutadans que mai han votat.
7. Votació que històricament ha tingut més participació.
8. Votació que històricament ha tingut més diferència percentual entre l'opció més votada i la menys.
9. Votació que històricament ha tingut menys diferència percentual entre l'opció més votada i la menys.
10. Número màxim de votacions en les que ha participat una mateixa persona i el país d'aquesta persona, tenint en compte tota la història de votacions d'Europa.

Les respostes del mòdul estadístic han de ser immediates i aquest ha d'estar sempre actualitzat amb la darrera informació de la BD, és a dir, NO es poden utilitzar *jobs* que s'executin per la nit i que emplenin les dades estadístiques (tot i que aquest sigui un procediment habitual en l'àmbit de les bases de dades per a un *datawarehouse*).

Requeriments de metodologia

L'estudiant haurà de dissenyar la BD i preparar els procediments emmagatzemats necessaris perquè es puguin realitzar les tasques anteriorment descrites. En cap cas ha d'implementar una interfície d'usuari amb un llenguatge de programació d'alt nivell.

Per realitzar el disseny de la BD es realitzarà primerament el model conceptual emprant el Diagrama E/R o UML incloent una llista amb totes les restriccions d'integritat que considerin rellevants. El consultor farà la funció de client i, per tant, qualsevol dubte que es presenti en quant a la realitat a descriure s'haurà de consultar amb ell. Un cop estudiat l'univers de discurs, caldrà crear els *scripts* de creació de taules, índexs, disparadors etc., i després s'implementaran els procediments que siguin necessaris.

Per a estandarditzar el sistema que s'ha de fer, es demana explícitament que els procediments emmagatzemats compleixin les condicions següents, que no podran ser obviades pel futur enginyer informàtic que desenvolupi el sistema:

- Com a mínim disposaran d'un paràmetre de sortida anomenat RSP, de tipus *string*, que indicarà si l'execució ha finalitzat amb èxit (valor 'OK') o si ha fracassat (valor 'ERROR+TIPUS D'ERROR')
- Disposaran de tractament d'excepcions.
- Emmagatzemaran totes les crides a procediments que es facin en una taula de *log*, emmagatzemant el procediment executat, els paràmetres d'entrada i els de sortida.

Els procediments emmagatzemats s'hauran d'especificar el millor possible, és a dir, han de tenir la suficient documentació associada per a que puguin ser utilitzats pels programadors de la capa de presentació, sense que aquestos hagin de mirar el codi dels procediments per saber que fan. Entre d'altres això implica:

- Descriure que fa el procediment a alt nivell
- Descriure els tipus i valors possibles de cada paràmetre d'entrada
- Descriure els tipus i valors possibles de cada paràmetre de sortida, incloent els diferents codis de d'error que pot retornar i el seu significat.

Adicionalment el codi dels procediments haurà de tenir comentaris que expliquin el seu comportament intern per a facilitar el manteniment posterior d'aquests.

L'estudiant, a part de presentar els *scripts* de creació de la BD, n'haurà de fer la inicialització amb un conjunt de dades suficient i presentar un joc exhaustiu de proves que permeti garantir el bon funcionament de les funcionalitats implementades, així com el control d'errors i de situacions d'excepció.

El sistema de gestió de BD emprat per a la part pràctica del treball serà Oracle. L'ús d'aquest sistema s'ha d'entendre com a valor afegit del treball, ja que actualment, gaudeix de molt bona posició a nivell de mercat.

Fonts d'informació

- Materials didàctics de les assignatures de Bases de dades I, Bases de dades II i Estructura de la informació.
- Scott Urman. "Oracle9i: Programación PL/SQL", Oracle Press Osborne, McGraw Hill, 2002. (ISBN: 978-84-481-3707-6).

Adicionalment, al pla docent de l'assignatura trobareu bibliografia complementària. En cas de consultar fonts addicionals, caldrà indicar-ho a bibliografia de la memòria del TFC.

Primera PAC

Un cop hagueu entès perfectament l'enunciat, haureu de realitzar la primera PAC, en aquesta haureu de presentar un pla de treball que ha d'incloure com a mínim:

- La descripció de que farà el sistema que dissenyareu i implementareu, explicant amb les vostres paraules quin serà l'abast total del projecte i els seus objectius (heu de demostrar que realment heu entès l'enunciat del TFC).
- La planificació detallada del projecte amb un diagrama de Gantt o similar.
- L'explicació aproximada de que penseu lliurar a la PAC2 i la PAC3 lligat a la planificació anterior.
- Els recursos humans, tècnics i materials que emprareu.
- Un anàlisi dels riscos del projecte i dels controls que aplicareu per mitigar-los per reduir el impacte negatiu que poden tenir.
- Una valoració inicial de les hores i del cost total del projecte.

Nota: en el món real aquesta primera PAC seria la que permetria triar entre un proveïdor tecnològic o un altre per realitzar el projecte, així que no s'ha de descuidar. A més, adaptant-la una mica, la podreu aprofitar com a primer capítol de la memòria.