

Arquitectura de Bases de Dades

PAC1: Bases de Dades Distribuïdes i Seguretat

Pregunta 1 (2.5 punts)

Enunciat

La base de dades GIMSport està formada per les relacions següents (claus primàries subratllades). A no ser que es digui el contrari, tots els atributs són obligatoris (no admeten valors nuls).

CUSTOMER(DNIC, Name, Surname, Age, Birth, eMail)
 MONITOR(DNIM, Name, Surname, Specialization, JoinDate)
 COURSES(Code, Name, Specialization, Requisites, Duration, Monitor)
 {Monitor} és clau forana que referencia {DNIM} MONITOR.
 ENROLLCOURSE(Customer, Course, Date) on
 {Customer} és clau forana que referencia {DNIC} CUSTOMER,
 {Course} és clau forana que referencia {Code} COURSE

La relació CUSTOMER conté informació dels clients del gimnàs, on l'edat és un camp numèric. La relació MONITOR conté la informació dels monitors que conformen la plantilla del gimnàs, on l'especialització pot ser "swimming", "fitness" o "ABS".

La relació COURSE conté informació sobre els cursos que s'imparteixen, així com el monitor encarregat de cada un d'ells. I per últim, la relació ENROLLCOURSE conté els clients que participen en cada un dels cursos oferts.

Es decideix aplicar l'estratègia de fragmentació següent:

CU1 = (Customer[DNIC, Name, Surname, Birth, Age])(Age>65)
 CU2 = (Customer[DNIC, Name, Surname, Birth, Age])(Age=65)
 CU3 = (Customer[DNIC, Name, Surname, Birth, Age])(Age<65)

M1 = Monitor(Specialization="Swimming")

M2 = Monitor(Specialization="ABS")

CR1 = Courses \bowtie M1

CR2 = Courses \bowtie M2

EC1 = EnrollCourse \bowtie CU1

EC2 = EnrollCourse \bowtie CU2

Es demana explicar l'estratègia de fragmentació descrita, així com avaluar el seu grau de correcció.

Criteris d'avaluació

Els criteris d'avaluació que s'aplicaran en la correcció d'aquesta pregunta són els següents:

- Es valorarà la qualitat de la resposta (concreció, no entrar en contradiccions etc.)
- Es valorarà la correcta identificació de les estratègies de fragmentació
- Es valorarà la correcta identificació dels defectes (si existeixen) de l'estratègia

Solució

S'ha dut a terme l'estratègia de fragmentació següent:

- **Customer: Fragmentació Híbrida:** CU1, CU2 i CU3 s'han construït aplicant en primera instància una fragmentació vertical sobre la taula *Customer*, considerant els atributs *DNIC*, *Name*, *Surname*, *Birth* i *Age* i a continuació, aplicant una fragmentació horitzontal en referència a l'edat dels clients.
- **Monitor: Fragmentació Horitzontal.** M1 i M2 s'han construït aplicant una estratègia de fragmentació horitzontal en referència a l'especialitat del monitor.
- **Courses: Fragmentació Horitzontal Derivada:** CR1 i CR2 s'han obtingut realitzant una operació de *semi join* entre la taula de cursos (*Courses*) i els fragments de la taula de monitors M1 i M2 que contenen els monitors amb especialitat *Swimming* i *ABS*.
- **EnrollCourse: Fragmentació Horitzontal Derivada:** EC1 i EC2 s'han obtingut realitzant una operació de *semi join* entre la taula de *EnrollCourse* i els fragments de la taula de clients CU1 i CU2 que contenen els clients majors de 65 anys i els que en tenen 65.

L'estratègia de fragmentació proposada no és correcta degut als aspectes següents:

- La **fragmentació híbrida** de la taula **Customer** no és correcta, ja que a nivell vertical s'obvia la informació corresponent al correu electrònic (*eMail*).
- La **fragmentació horitzontal** de la taula **Monitor** no és correcta ja que obtenim dos subconjunts de la taula Monitor, els d'especialitat *Swimming* i *ABS*, però perdem els d'especialització *fitness*.
- La **fragmentació horitzontal derivada** de la taula **Courses** no és correcta ja que hereta la problemàtica descrita al punt anterior en relació a la fragmentació de la taula **Monitor**.
- La **fragmentació horitzontal derivada** de la taula **EnrollCourse** no és correcta. Aquesta derivació no és completa ja que s'obvia CU3 que conté els clients menors de 65 anys.

Pregunta 2 (2.5 punts)

Enunciat

El gimnàs que farà ús del model de dades descrit en l'exercici 1, molt preocupat pel compliment del RGPD (Reglament General de Protecció de Dades), d'obligat compliment des del 25 de maig de 2018, està revisant les actuacions a realitzar abans del canvi de l'aplicatiu que gestionarà les dades dels clients. L'equip tècnic responsable del sistema d'informació es planteja les qüestions següents:

- Quins són els aspectes clau que cal tenir en compte en el formulari que es publicarà a la web per tal que la gent pugui sol·licitar informació sobre el gimnàs i els serveis que ofereix, i que permetrà després realitzar campanyes de màrqueting d'enviament de correus electrònics?
- Quins són els principis bàsics relatius al tractament de dades personals?
- Quins drets als ciutadans s'incorporen en l'entrada en vigor del RGPD, que no es trobaven en la LOPD (Llei Orgànica de Protecció de Dades)? Explica breument el significat d'aquests.

Nota: la resposta a tot aquest exercici no hauria de superar l'extensió d'1 pàgina i mitja, es demana incloure les fonts consultades.

Criteris d'avaluació

Els criteris d'avaluació que s'aplicaran en la correcció d'aquesta pregunta són els següents:

- La distribució dels punts és d'un punt per les preguntes *b* i *c*, i 0.5 per la pregunta *a*.
- Es valorarà la qualitat de la resposta, així com la concreció en l'argumentació d'aquesta
- Es valorarà la inclusió de les fonts consultades per donar resposta a la pregunta
- Es valorarà no superar l'extensió indicada per la resposta
- Les preguntes no contestades no penalitzen

Solució

- L'RGPD requereix que l'interessat doni consentiment mitjançant una acció clara afirmativa, per tant les caselles ja marcades o el consentiment tàcit ja no es considera vàlid. L'objectiu és que l'interessat sigui el que indiqui que vol rebre informació del gimnàs, i per tant, en el formulari que s'ha dissenyat es mostrarà clarament un camp a tal efecte.

Aquest camp farà referència a la informació de la política de privacitat i especificarà que aquest accepta les condicions d'aquesta, aquesta casella no pot estar marcada d'inici, sinó que l'usuari ha de marcar-la.

Cal revisar també l'avís legal de la web indicant clarament la següent informació:

- Identitat del responsable de gestió de les dades.
- Dades que s'estan recollint i utilitzant.
- Categoria de les dades recollides.
- Finalitat d'ús de les dades.
- Període de conservació de les dades personals recollides.
- Si es preveu fer cessió d'aquestes dades.
- Com exercir els seus drets.

Font: <https://www.aepd.es>

b) A continuació s'identifiquen els principis referents al tractament de les dades personals:

- Principi de licitud, lleialtat i transparència, el reglament determina que els tractaments de les dades personals han de ser lleials, i les empreses han de ser transparents en referència al tractament de les dades, i informar a l'interessat de forma oberta i transparent.
- Recollides amb finalitats determinades, explícites i legítimes, el tractament de les dades ha d'estar limitat a les finalitats per les quals es van recollir originàriament i aquestes han de ser legítimes.
- Minimització de dades, durant la recollida de dades, sols s'han de sol·licitar les absolutament necessàries per a la finalitat.
- Limitació del període de conservació, les dades personals han de ser conservades només mentre siguin necessàries, és a dir, s'han d'eliminar en el moment que s'hagi acomplert la finalitat per la qual s'havien recollit.
- Integritat i confidencialitat, les dades personals han de ser tractades de forma que es pugui garantir la seguretat apropiada, incloent la protecció contra el tractament no autoritzat. A part, cal garantir que les dades no poden ser modificades per persones no autoritzades.
- Exactes, les dades personals dels interessats han de ser sempre precises i estar actualitzades, això implica que aquestes s'han d'actualitzar quan l'interessat ho sol·liciti.

c) Fins a l'entrada en vigor del Reglament, els drets garantits en el control de les dades personals eren els anomenats drets ARCO:

- Dret d'accés
- Dret de rectificació
- Dret de cancel·lació
- Dret d'oposició

Aquests s'amplien amb:

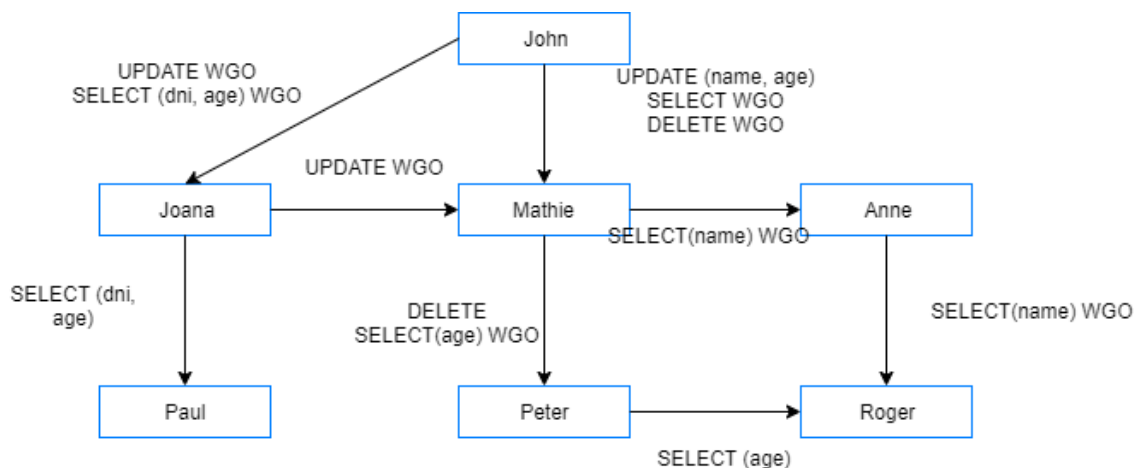
- Dret a l'oblit: Dret del ciutadà a que les seves dades se suprimeixin quan ja no són necessàries per a la finalitat per la qual es van recollir, o quan es revoca el consentiment en el qual es basava el tractament, etc.
- Dret a la limitació de tractament: Dret del ciutadà a que sota petició d'aquest no s'apliquin a les seves dades les operacions de tractament que en cada cas correspondrien.
- Dret a la portabilitat de les dades: Complement del dret d'accés. Permet obtenir les dades proporcionades a una organització o a transmetre-les de forma directa a una altra.

També cal fer notar que el dret de cancel·lació passa a ser anomenat dret de supressió.

Pregunta 3 (2.5 punts)

Enunciat

Per tal de realitzar aquest exercici, ens basarem en el model de dades descrit a l'exercici 1 d'aquesta mateixa PAC. Donat el graf següent d'autoritzacions sobre la taula CUSTOMER, propietat de l'usuari John.



Es demana:

- Proposar una seqüència de sentències d'autoritzacions, en SQL estàndard, indicant l'usuari que executa les sentències, que ens condueixi al graf d'autoritzacions anterior. Tingueu en compte que cal resoldre el problema des d'un punt de vista teòric, sense executar-lo a nivell pràctic sobre Oracle i assumint que, inicialment, els usuaris només tenen el privilegi CONNECT sobre la base de dades, excepte l'usuari John propietari de la taula i amb tots els permisos garantits.
- Tornar a dibuixar el graf d'autoritzacions després de que l'usuari John executi la sentència següent:

REVOKE SELECT ON CUSTOMER FROM "Mathie" CASCADE

- Tornar a dibuixar el graf d'autoritzacions, partint de la situació inicial, després de que s'executin les següents sentències de forma seqüencial, descrivint l'efecte de cadascuna d'elles sobre els privilegis dels usuaris:

Usuari Paul:

- GRANT SELECT(dni) ON CUSTOMER TO "Peter" WITH GRANT OPTION

Usuari Mathie:

2) GRANT SELECT(email) ON CUSTOMER TO "Paul"

Usuari Joana:

3) GRANT SELECT(dni,age) ON CUSTOMER TO "Paul" WITH GRANT OPTION

Usuari Ane:

4) REVOKE SELECT(email) ON CUSTOMER FROM "Roger" CASCADE

Criteris d'avaluació

Els criteris d'avaluació que s'aplicaran en la correcció d'aquesta pregunta són els següents:

- La distribució dels punts és de 0.75 per les preguntes *b i c*, i 1 punt per la pregunta *a*.
- Es valorarà la qualitat de la resposta
- Es valorarà la coherència entre les sentències i el graf
- Les preguntes no contestades no penalitzen

Solució

a) Les instruccions a executar serien les següents:

Usuari John:

```
GRANT SELECT ON CUSTOMER TO "Mathie" WITH GRANT OPTION
GRANT DELETE ON CUSTOMER TO "Mathie" WITH GRANT OPTION
GRANT UPDATE(name, age) ON CUSTOMER TO "Mathie"
GRANT UPDATE ON CUSTOMER TO "Joana" WITH GRANT OPTION
GRANT SELECT(dni, age) ON CUSTOMER TO "Joana" WITH GRANT OPTION
```

Usuari Joana:

```
GRANT SELECT(dni, age) ON CUSTOMER TO "Paul"
GRANT UPDATE ON CUSTOMER TO "Mathie" WITH GRANT OPTION
```

Usuari Mathie:

```
GRANT DELETE ON CUSTOMER TO "Peter"
GRANT SELECT (age) ON CUSTOMER TO "Peter" WITH GRANT OPTION
GRANT SELECT (name) ON CUSTOMER TO "Anne" WITH GRANT OPTION
```

Usuari Peter:

```
GRANT SELECT(age) ON CUSTOMER TO "Roger"
```

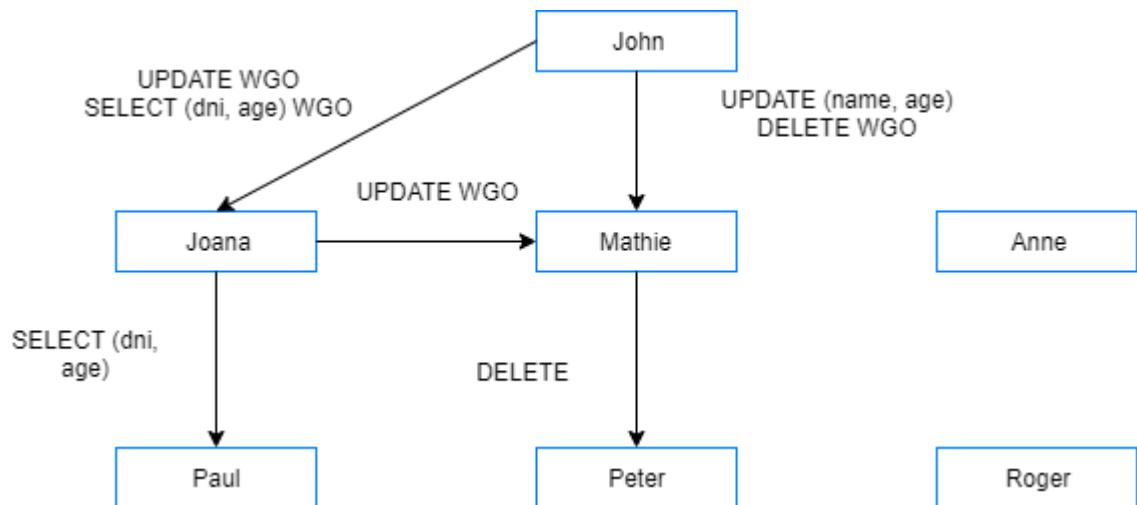
Usuari Anne:

```
GRANT SELECT(name) ON CUSTOMER TO "Roger" WITH GRANT OPTION
```

Usuaris Roger i Paul:

No efectuen cap assignació de privilegis cap a altres usuaris, només en són receptors.

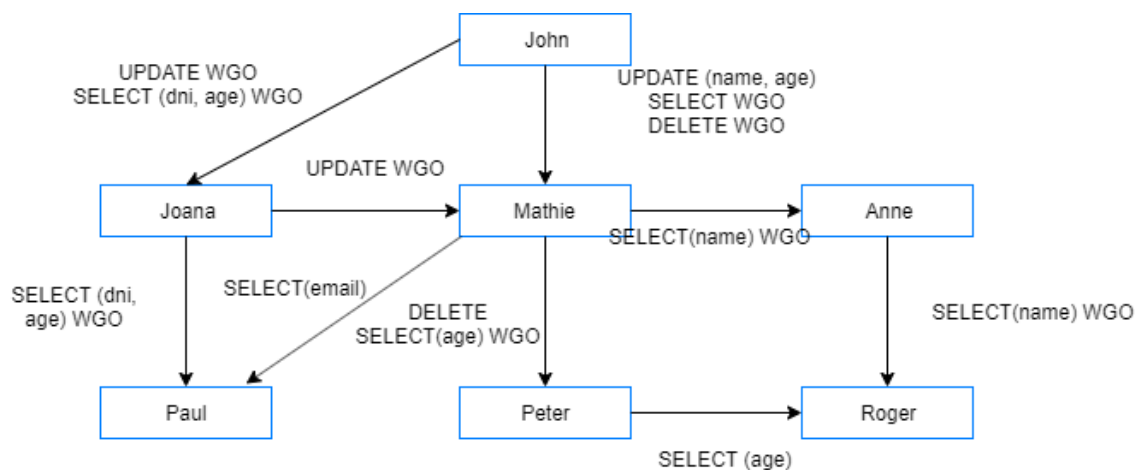
b) El graf quedaria de la següent manera:



c) A continuació es descriuen els efectes de cadascuna de les sentències:

- 1) La sentència donarà error ja que l'usuari "Paul" no té privilegis per assignar aquest permís, no se li havia atorgat amb WITH GRANT OPTION.
- 2) La sentència s'executarà correctament i afegirà en el privilegi de SELECT sobre el camp mail.
- 3) La sentència s'executarà correctament proporcionant privilegis de SELECT(dni,age) WGO a l'usuari "Paul".
- 4) La sentència donarà error ja que no es pot revocar un permís que no disposa l'usuari.

El graf quedaria de la manera següent:



Pregunta 4 (2.5 punts)

Enunciat

Com arquitecte de bases de dades de l'empresa HotelRooms, dedicada a la gestió d'hotels, se't demana solucionar els problemes actuals de rendiment dels aplicatius que permeten gestionar la informació dels hotels.

A continuació es detallen les diferents relacions que formen part de la base de dades (claus primàries subratllades):

Guest (id, name, address, phone, email, birth_date)

Employee_ht (id, name, address, phone, email, hotel)

{hotel} és clau forana que referencia {id} Hotel

Hotel (id, name, category, address, location)

{location} és clau forana que referencia {id} Location

Location (id, name, state)

BookingRoom(cod, id_hotel, id_guest, date_in, date_out)

{id_hotel} és clau forana que referencia {id} Hotel

{id_guest} és clau forana que referencia {id} Guest

L'empresa té una seu central on trobem els departaments corresponents a la gestió de personal i a la gestió comercial. Els dos departaments estan físicament ubicats en el mateix edifici. L'empresa, tal i com s'ha comentat, gestiona diferents hotels situats en diferents localitzacions, i clients que s'allotgen en aquests hotels.

Els requisits de dades són els següents:

- El departament de gestió de personal s'encarrega de mantenir les dades dels treballadors, els hotels poden consultar el *id*, *name*, *phone* dels seus treballadors, però no els de la resta.
- La informació dels hostes dels hotels és gestionada pels diferents hotels, però compartida per tots ells.

- El departament de gestió comercial necessita consultar les dades dels hostes, concretament el nom, email i data de naixement per fer promocions, i no té accés a cap més informació d'ells. També pot consultar la informació de reserves fetes per cada hoste.
- Cada hotel gestiona les reserves pròpies, però només necessita consultar la informació de les reserves del propi hotel. No es permet la consulta de reserves de la resta d'hotels.
- Cada hotel s'encarrega de mantenir la informació pròpia de l'hotel i no necessita accedir a cap informació dels altres hotels que formen part de la companyia. El departament de gestió comercial consulta aquesta informació.
- El departament comercial necessita consultar la informació corresponent a les localitzacions dels hotels per determinar el tipus de publicitat que es realitza en cada una d'aquestes.

El departament de tecnologia de l'empresa, ha determinat on cal situar els diferents de servidors per distribuir les dades de la base de dades i millorar el rendiment dels aplicatius, en aquests moments trobem que:

- Cada hotel disposa d'un servidor de BD.
- El departament de gestió comercial i de gestió de personal disposen d'un servidor de BD cada un, localitzats al CPD de la seu central.

A partir dels requeriments de dades indicats es demana proposar una estratègia de distribució de dades per a les diferents relacions. Expliqueu breument, per a cada relació de l'esquema de la base de dades, l'opció de distribució escollida ja sigui fragmentació, especificant quin tipus s'utilitzaria, o replicació per a cada entitat del sistema.

Criteris d'avaluació

Els criteris d'avaluació que s'aplicaran en la correcció d'aquesta pregunta són els següents:

- Es valorarà l'estratègia de fragmentació i de replicació proposada.
- Es valorarà la qualitat de la resposta.
- Es valorarà la concreció en l'argumentació de la resposta.

Solució

Les necessitats d'informació i dades descrites en l'enunciat es poden resumir en:

- **Guest:** Els diferents hotels s'encarreguen de mantenir les dades dels hostes, donat que aquesta informació es compartida per a tots els hotels, quan es realitza qualsevol operació primer es guarda en el servidor local i a continuació es distribueix a la resta de servidors per a que estigui disponible a tots els hotels.

Les dades dels hostes es replicaran als diferents servidors dels hotels.

El departament comercial necessita conèixer la informació corresponent a *name*, *mail* i *birth_date*, per això es determina realitzar una **fragmentació vertical** per aquests camps, incloent l'identificador (*id*) i després de fragmentar les dades es replicaran al servidor del departament de gestió comercial.

- **Employee_ht:** En el cas de la informació dels treballadors, fragmentarem la taula per hotel, així garantirem que només poden accedir a la seva informació, a part cal tenir en compte que l'hotel només accedeix al *id*, *name* i *phone*, la resta d'informació la gestiona el departament de personal. Ens trobem davant d'una **fragmentació híbrida**, horitzontal per hotel i vertical pels atributis indicats anteriorment.

Les dades fragmentades, per hotel, es repliquen després a la base de dades del departament de personal.

- **Hotel:** Aquesta informació es mantindrà centralitzada en el servidor del departament comercial, i els diferents hotels accediran aquí per mantenir les dades. Cal tenir en compte que es tracta d'una informació que varia poc i per tant, no és problema ubicar-la en aquest servidor.
- **Location:** Aquesta informació es mantindrà centralitzada en el servidor del departament comercial, ja que aquest és qui fa ús d'aquestes dades.
- **BookingRoom:** En el cas de la les reserves, fragmentarem la taula per hotel, així garantirem que només poden accedir a la seva informació. En aquest cas farem una **fragmentació horitzontal** per hotel.

Les dades fragmentades, per hotel, es repliquen després a la base de dades del departament comercial.

Recursos

Per tal de resoldre aquesta PAC caldrà utilitzar els recursos que s'enumeren a continuació:

- Mòduls 3 (bases de dades distribuïdes) i 4 (seguretat) de l'assignatura Arquitectura de Bases de Dades.

Criteris de valoració

A l'enunciat de cada exercici s'indica el valor del mateix sobre la puntuació total de la PAC. Aquesta activitat representa el 50% de la nota d'avaluació continuada de l'assignatura i la seva realització és **obligatòria** per tal de superar l'assignatura.

No s'acceptaran ni es tindran en compte els lliuraments realitzats fora dels terminis indicats al calendari de l'aula.

Format i data de lliurament

Heu d'enviar la PAC1 a la bústia de lliuraments com un arxiu .pdf el nom del qual contingui el codi de l'assignatura, el vostre cognom i el vostre nom, així com la indicació de que es tracta de la PAC1. La resolució i lliurament d'aquesta activitat **només és pot fer de forma individual**.

La data límit per lliurar la PAC1 és el **01/12/2019**.