

# Conceptes bàsics y models NoSQL PAC1

## Exercici 1 (30%)

Realitza la lectura del següent article: [Persisting big-data: The NoSQL landscape.](#)

Es demana respondre de manera concisa (una pàgina com a màxim per a la resposta de les 5 preguntes juntes) a les següents preguntes:

1. Per què algunes bases de dades NoSQL adopten el model BASE?
2. Què va demostrar Seth Gilbert i Nancy Lynch sobre el teorema de Brewer?
3. Si  $N$  representa el nombre de nodes on una clau es troba replicada,  $W$  el nombre de nodes necessaris per considerar que una escriptura s'ha realitzat amb èxit i  $R$  el nombre de nodes on es realitza una lectura. Suposem que es disposa d'un sistema distribuït on  $N = 5$ ,  $W = 4$  i  $R = 5$ , què es pot dir d'aquest sistema en termes de consistència i sobre les seves lectures i escriptures?
4. Totes les bases de dades comentades en el text realitzen la seva persistència sobre disc? En cas que no sigui així, explica en quins casos no ho fan.
5. Explica breument quins són els principals sistemes de redistribució de claus en una base de dades de tipus clau-valor.

## Exercici 2 (20%)

A partir de la lectura del llibre *NoSQL Distilled* indiqueu què us semblen les següents afirmacions.

### Afirmació 1

Un model que ignora les agregacions permet fàcilment veure les dades des de diferents perspectives, de manera que serà una bona elecció quan no hi ha una estructura clara de manipulació de les dades.

### Afirmació 2

Ja que una de les característiques de les bases de dades NoSQL és l'ús d'esquemes implícits (*schemaless*), enfront dels esquemes fixos que fan servir les bases de dades relacionals, podríem afirmar que l'ús d'esquemes implícits evita situacions complicades

quan diverses aplicacions desenvolupades per diferents persones / equips accedeixen a la mateixa base de dades.

### Afirmació 3

Una base de dades relacional pot tractar amb dades no uniformes (és a dir cada fila de la taula podria tenir diferents camps d'informació) sense que es produeixin cap tipus d'anomalia en les taules.

### Afirmació 4

El terme de "persistència políglota" indica que en comptes d'utilitzar un únic tipus de base de dades en un projecte, el que s'ha de fer és combinar diferents tipus de bases de dades d'acord amb les diferents tipus de circumstàncies (naturalesa de les dades i manipulació dels mateixos).

Per a cadascuna de les afirmacions:

- Indiqueu si és certa o falsa. **No seran vàlides les respostes que no indiquin si l'afirmació és certa o és falsa.**
- Justificar breument la vostra resposta fent referència al llibre bàsic de l'assignatura o els apunts.

Es valorarà la concisió (una pàgina com a màxim per les 4 afirmacions juntes).

## Exercici 3 (30%)

Considera les taules d'un sistema relacional per gestionar llibres d'una biblioteca:

### Taula Autors

<u>Columna</u>	<u>Tipus</u>	<u>NULS</u>	<u>DEFECTE</u>	<u>Clau-Pri</u>	<u>Únic</u>
Autor_ID	INT	No		Sí	Sí
Nom	VARCHAR(10)	No			
Cognoms	VARCHAR(35)	No			
Data_Naixament	Date	No			
Data_Mort	Date	Sí	NULL		

### Taula Llibres

<u>Columna</u>	<u>Tipus</u>	<u>NULS</u>	<u>DEFECTE</u>	<u>Clau-Pri</u>	<u>Únic</u>
ISBN	INT	No		Sí	Sí
Títol	VARCHAR(35)	No			
Autor_ID	INT	No			No
Editorial	VARCHAR(20)	No	"		
Suport	VARCHAR(35)	No	"		
Data de registre	DATE	Sí	NULL		
País	VARCHAR(20)	Sí	NULL		

### Taula Usuaris

<u>Columna</u>	<u>Tipus</u>	<u>NULS</u>	<u>DEFECTE</u>	<u>Clau-Pri</u>	<u>Únic</u>
Nom	VARCHAR(10)	No			
Cognoms	VARCHAR(35)	No			
DNI	VARCHAR(15)	No		Sí	Sí
Email	VARCHAR(35)	No			Sí
Telèfon	INT	Sí	NULL		

### Taula Prèstecs

<u>Columna</u>	<u>Tipus</u>	<u>NULS</u>	<u>DEFECTE</u>	<u>Clau-Pri</u>	<u>Únic</u>
registre	INT	No		Sí	Sí
Id_Llibre	INT	No			No
Id_Usuari	VARCHAR(15)	No			
Data_prèstec	DATE	No			
Data_devolució	DATE	Sí			

Després de realitzar un estudi sobre les consultes que es volen fer, s'ha arribat a la conclusió que es volen només tres tipus de consultes:

1. De cada usuari es vol conèixer el nom i DNI, i un llistat d'activitat d'usuari per mesos. A cada mes es desitja conèixer les dades de cada llibre que està o ha estat en préstec durant aquest mes: ISBN, títol, Autor\_ID, Editorial i Data devolució.
2. De cada editorial es vol conèixer nom de l'editorial i un llistat per mesos de préstec. A cada mes es desitja conèixer les dades de cada llibre pertanyent a l'editorial que està o ha estat en préstec durant aquest mes: ISBN, títol, Autor\_ID, DNI de l'usuari al que se li va prestar i data de préstec.
3. De cada autor (identificador, nom i cognom) es vol conèixer un llistat de préstecs agrupats per mesos. A cada mes es vol conèixer de cada llibre de l'autor: el títol del llibre, el nombre de préstecs del llibre, l'editorial a la qual pertany cada llibre, i un llistat dels DNIs de les persones que van sol·licitar el préstec dels exemplars.

Un cop analitzades el tipus de consultes que es volen dur a terme, s'ha optat per usar un model de dades orientat a agregats. Es demana definir els agregats necessaris que satisfacin i optimitzin les consultes plantejades. Concretament per cada agregat s'ha d'indicar:

- Breu explicació de l'estructura de l'agregat.
- Representació gràfica de l'agregat proposat usant un model gràfic de caixes niuades com el que apareix a la Figura 1.1 del Capítol 1 del llibre bàsic de l'assignatura.

## Exercici 4 (20%)

Es demana buscar un cas d'aplicació d'una base de dades NoSQL i contestar les següents preguntes (màxim 2 pàgines per a les 4 preguntes juntes):

1. Descriu el problema de persistència de dades que s'ha resolt.
2. Justifica les raons per les quals no és recomanable l'ús d'una base de dades relacional.
3. Justifica les raons per les quals és recomanable la base de dades que s'ha utilitzat com a solució.
4. Indicar les referències utilitzades per desenvolupar l'exercici.

No es pot reproduir de forma textual cap frase de les referències utilitzades. A continuació es mostren alguns exemples. **Aquests exemples no es poden utilitzar per resoldre l'exercici (si es fan servir algun d'aquests exemples la pregunta estarà anul·lada), s'han de buscar-ne altres de diferents:**

- Ús d'un [model clau valor per a la gestió en el àmbit de la salut](#).
- Ús d'un [model en graf per analitzar els papers de Panamà](#).
- Ús d'un [model documental per oferir un servei de continguts personalitzat i en temps real](#).
- Ús d'un [model orientat a columnes per a un sistema de moocs](#).

## Criteris de valoració

Els apartats 1) i 3) tenen un pes del 30%, i els apartats 2) i 4) tenen un pes del 20% cadascun. Es valorarà, per a cada apartat, la validesa de la solució i la claredat de l'argumentació.

## Format i data de lliurament

Heu d'enviar la PAC a la bústia de Lliurament i registre d'AC disponible a l'aula (apartat Avaluació). El format de l'arxiu que conté la vostra solució pot ser .pdf, .odt, .doc i .docx. Es recomana el format .pdf. Per a altres opcions, si us plau, contacteu prèviament amb el vostre professor col·laborador. El nom del fitxer ha de contenir el codi de l'assignatura, el vostre cognom i el vostre nom, així com el nom de l'activitat (PAC1). Per exemple nomcognom1\_nosql\_pac1.pdf

La data límit per lliurar la PAC1 és el **24 de març**.

## Propietat intel·lectual

En presentar una pràctica o PAC que faci ús de recursos aliens, s'ha de presentar juntament amb ella un document en què es detallin tots ells, especificant el nom de cada recurs, el seu autor, el lloc on es va obtenir i el seu estatus legal: si l'obra està protegida pel copyright o s'acull a alguna altra llicència d'ús (Creative Commons, llicència GNU, GPL etc.). L'estudiant haurà de assegurar que la llicència que sigui no impedeix específicament seu ús en el marc de la pràctica o PAC. En cas de no trobar la informació corresponent haurà d'assumir que l'obra està protegida pel copyright.

Caldrà, a més, adjuntar els fitxers originals quan les obres utilitzades siguin digitals, i el seu codi font, si així correspon.