Rapport TP BDD

BLANC Loriane

GUIRAUD Vincent ZENNARO Thomas

HELLOUIN DE CENIVAL Zacharie

BOURLOT Xavier

3IMACS-AE-C, Binôme 8

$3~\mathrm{juin}~2019$

Table des matières

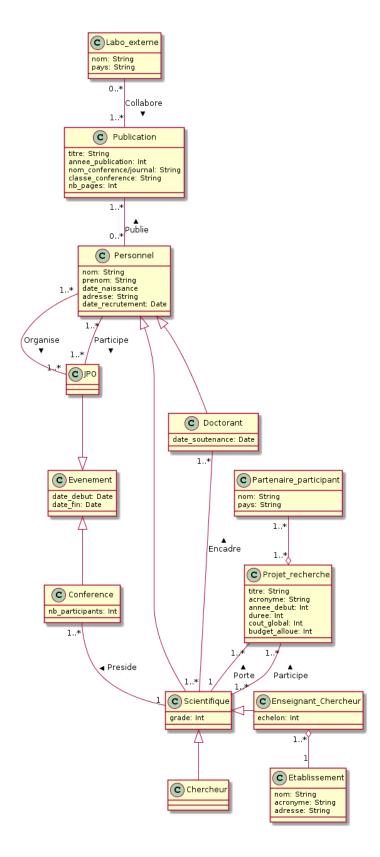
1	Introduction	2
2	Conception UML	2
3	Modèle Relationnel	3
4	Normalisation	4
5	Création des tables	4
6	Remplissage des tables	4
7	Écriture des Requêtes 7.1 Algèbre relationnelle	4
8	Conclusion	7
	Annexes A.1 Requêtes SQL	8

1 Introduction

Nous souhaitons mettre en place une base de données afin de gérer les évaluations du laboratoire du LAAS-CNRS. Pour cela nous commencerons par concevoir un diagramme de classe UML, puis nous chercherons à traduire ce diagramme en modèle relationnel en utilisant SQL. L'étape suivante consiste à remplir la base de données avec des valeurs inventées afin de pouvoir tester la viabilité de notre modèle par le biais de différentes requêtes.

2 Conception UML

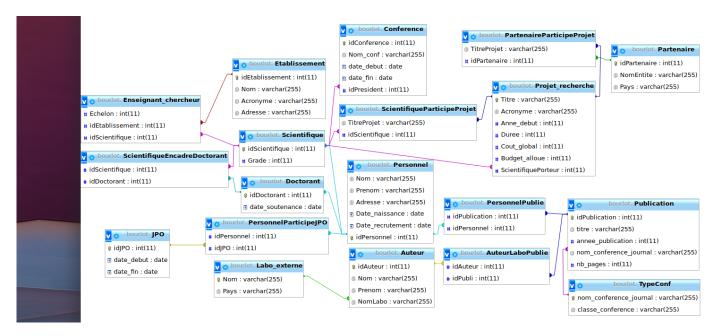
A partir du cahier des charges qui nous a été fourni, nous avons conçu un diagramme de classe. Nous avons bien pris en compte le fait qu'il y a deux types de personnel : les scientifiques et les doctorants, et qu'un scientifique peut encadrer un doctorant. Nous avons procédé de même pour les autres classe, en suivant à la lettre le cahier des charges qui nous as permis de construire les classes, définir leurs arguments, et les liens entre eux ainsi que leur cardinalité. Nous avons ensuite traduit ce digramme de classe en UML. Voici le diagramme ainsi obtenu :



3 Modèle Relationnel

Afin de pouvoir implémenter notre base de données, nous avons utilisé le langage SQL pour créer nos tables que nous avons pu visualiser sur phpMyAdmin. Nous avons ajouté une classe Auteur et une classe AuteurPublie pour pouvoir ajouter des collaborateurs externes dans les publications. De même, nous avons ajouté une classe d'association entre Scientifiques et Doctorants, JPO et Personnel, Scientifique et Projet de recherche, et Partenaire avec Projet de recherche.

Ces classes d'association permettent de faire le lien entre deux classes pour lesquelles une action est menée (participation, encadrement de personnel, publication etc.). Nous avons donc créé le modèle relationnel suivant :



4 Normalisation

Nous n'avons pas rencontré de problèmes particuliers concernant la normalisation, toutes les formes normales semblaient être respectées. Nous pouvons donc passer à l'étape suivante : la création des tables.

5 Création des tables

Dans cette étap, nous n'avons eu qu'à créer les tables à partir des modèles précédemment établis. La syntaxe étant la même pour la création de toutes les tables, la difficulté consistait à ne pas oublier d'arguments et leur affecter le bon type. Il fallait également veiller à la présence des clés étrangères qui lie certaines classes, notamment les classes d'association. Nous avons donc créé les tables dans le langage SQL.

6 Remplissage des tables

Le remplissage des tables s'est fait de manière à pouvoir tester les différentes requêtes; il était en effet nécessaire d'avoir des données correspondant aux interrogations SQL de la base de donnée. Nous avons essayé de faire en sorte que la BDD soit la plus complète possible, mais nous avons dû modifier plusieurs fois l'ensemble des tables afin de correspondre à une requête précise que nous n'avions pas encore traiter, ce qui a été un travail fastidieux et peu intéressant. Le code que nous avons écrit respecte donc ces contraintes et permet de tester les requêtes de la prochaine partie.

7 Écriture des Requêtes

7.1 Algèbre relationnelle

Cette partie fut la plus chronophage. En effet, la traduction des différentes requêtes en algèbre relationnelle, plus formelle, s'est révélée délicate à la fois en raison de la complexité de la base de données, et parfois à cause des différentes interprétations possibles de la requête. Nous nous sommes donc répartis les requêtes, en s'entraidant en cas de difficulté, car certaines requêtes étaient bien plus fastidieuses que d'autres. La cohérence entre les requêtes et l'architecture en évolution continue de notre base de données était également un obstacle que nous avons dû surmonter.

7.2 SQL

Pour plus de lisibilité nous n'avons pas mis l'ensemble des requêtes, le but étant d'en présenter quelques unes pour visualiser le fonctionnement. L'intégralité des requêtes est fournie en annexe.

REQUETE 1 : L'idée ici était de d'abord récupérer le nom et l'id des scientifiques encadrant le doctorant (on a choisi le doctorand dont l'id était 1). Pour obtenir le grade et l'id, nous avons dû passer par le biais de la table "Scientifique" et procéder demanière identique qu'avec le nom, qui lui était situé dans "Personnel". Il a ensuite suffit de joindre ces deux tables afin de récupérer juste le nom, associé au grade, les id récupérés permettant d'assurer une bonne correspondance.

Nous obtenons le résultat suivant :



REQUETE 6 : Nous utilisons la fonction "COUNT (*)" puis GROUP BY à partir de la table Publication afin de pouvoir les compter.

```
--6. Le nombre de publications par annee de tout le laboratoire

(SELECT annee_publication, COUNT(*) as cntPubli FROM Publication CROUP BY annee_publication);
```

Nous obtenons le résultat suivant :



REQUETE 16: L'élaboration de cette requête est assez évidente, car il a suffit d'exclure de la liste des scientifiques tous ceux qui étaient présents dans "ScientifiqueEncadreDoctorant" ou "ScientifiqueParticipeProjet" ou "PersonnelPublie". Pour récupérer le nom et prénom, il a ensuite fallu passer par le biais de la table "Personnel".

Nous obtenons le résultat suivant :

```
nom prenom
Kine Miss
```

REQUETE 21 : Pour la requête 21 (trouver le nombre de collaborateurs par pays), nous avons joint les tables Auteur et Labo externe qui ont leur nom de laboratoire en commun (ce qui veut dire qu'ils ont collaboré), ensuite nous avons groupé ces laboratoires par pays, et enfin, nous avons compté le nombre de laboratoire qui collabore par pays. Voici son code en SQL :

```
--21. Le nombre de collaborateurs par pays

SELECT pn. Pays, COUNT(idAuteur)

FROM(SELECT Pays, Nom FROM Labo_externe) as pn

LEFT JOIN(SELECT idAuteur, NomLabo FROM Auteur) as a

ON pn. Nom=a. NomLabo

GROUP BY pn. Pays
```

```
pays count
France 3
URSS 1
Estonie 1
```

REQUETE 26 : Nous comparons le nombre de doctorants recrutés chaque année par un scientifique avec le nombre d'années écoulées entre le recrutement du scientifique et la date actuelle.

```
Les scientifiques qui recrutent au moins un doctorant par annee
     -26.
2
   SELECT da.idScientifique FROM
   (SELECT sd.idScientifique, COUNT(sd.DateDoc) as NbDocs, AVG(EXTRACT(YEAR FROM p.Date_recrutement)) as
       DateSci
   FROM (SELECT idPersonnel, Date recrutement FROM Personnel) as p
   JOIN (SELECT DISTINCT s.idScientifique, EXTRACT(YEAR FROM d. Date recrutement) as DateDoc
   FROM
   (SELECT id Scientifique, id Doctorant FROM Scientifique Encadre Doctorant) as s
10
11
   JOIN
   (SELECT idPersonnel, Date_recrutement FROM Personnel WHERE idPersonnel IN (SELECT idDoctorant FROM
12
       Doctorant)) as d
   ON s.idDoctorant=d.idPersonnel ) as sd
   ON sd.idScientifique=p.idPersonnel
14
   GROUP BY sd.idScientifique) as da
15
   WHERE (da.nbdocs>= (EXTRACT(YEAR FROM CURRENT DATE) - da.DateSci))
```

Nous obtenons le résultat suivant :

idscientifique 6

REQUETE 30 : Pour trouver le nombre de conférences de classe A organisées par le Laas par année, nous avons joint les tables Conférence et TypeConf pour lesquelles le nom de conférence était le même, avec uniquement des classes A (ce qui permet d'obtenir la liste des conférences de classe A données par le Laas), puis nous les avons regroupées par année et nous avons compté le nombre de conférence correspondant à l'aide d'un "COUNT(nom conférence)", tout en gardant uniquement les attributs année et nombre de conférences à l'aide d'un Select, puisque ce sont les informations qui nous intéressent. Voici le code en SQL :

annee	nbconf
2019	1
2018	2
2001	1
2000	1

8 Conclusion

Ce projet nous a permis d'aborder l'ensemble des notions vues en cours, en TD et de les mettre en pratique. Mise à part la requête numéro 24, nous les avons toutes testées avec notre base de données et nous avons pu les valider. Le cheminement nécessaire pour aboutir à une base de données fiable et fonctionnelle nous a permis de prendre conscience des écueils à éviter au cours d'un projet de cette ampleur. Il est par exemple très important de ne pas hésiter à modifier la structure du projet tout au long de la conception, même au moment de l'élaboration des requêtes.

A Annexes

A.1 Requêtes SQL

Ensemble de nos requêtes traduites en SQL ainsi que le résultat obtenu pour chacune :

A.1.1 R1

```
--1. Le nom et les grades des encadrants d un doctorant donne

SELECT p.Nom, s.grade FROM

(SELECT Nom, idPersonnel FROM Personnel WHERE idPersonnel IN (SELECT idScientifique FROM ScientifiqueEncadreDoctorant WHERE idDoctorant=1)) as p

JOIN

(SELECT s.grade, s.idScientifique FROM Scientifique s WHERE s.idScientifique IN (SELECT idScientifique FROM ScientifiqueEncadreDoctorant WHERE idDoctorant=1)) as s

ON s.idScientifique=p.idPersonnel
```

Nous obtenons le résultat suivant :



A.1.2 R2

```
--2. Les pays avec qui un scientifique donne collabore

SELECT Pays FROM Labo_externe WHERE Nom IN(SELECT NomLabo FROM Auteur WHERE idAuteur IN(SELECT idAuteur FROM AuteurLaboPublie WHERE idPubli IN(SELECT idPublication FROM PersonnelPublie WHERE idPersonnel =8)));
```

Nous obtenons le résultat suivant :



A.1.3 R3

```
-3. Les noms et les pays des auteurs collaborateurs d'un scientifique donne en 2018

SELECT nnl.Nom, l.Pays FROM
(SELECT Nom, NomLabo FROM Auteur WHERE idAuteur IN(SELECT idAuteur FROM AuteurLaboPublie WHERE idPubli IN(SELECT idPublication FROM PersonnelPublie WHERE idPersonnel=8)

INTERSECT
(SELECT idPublication FROM Publication WHERE annee_publication=2018))) as nnl

JOIN
(SELECT Nom, Pays FROM Labo_externe) as l
ON nnl.NomLabo=l.Nom
```

Nous obtenons le résultat suivant :



A.1.4 R4

```
-4. Le nombre de collaborateurs d'un scientifique donne en 2018

SELECT COUNT(idPersonnel) as NbCollab FROM PersonnelPublie
WHERE (idPublication IN

(SELECT idPublication FROM Publication
WHERE annee_publication=2018 AND idPublication IN

(SELECT idPublication FROM PersonnelPublie
WHERE idPersonnel=10)

) AND idPersonnel!=10)
```

NbCollab 2

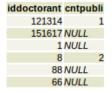
A.1.5 R5

```
--5. Pour chaque doctorant, on souhaiterait recuperer le nombre de ses publications

SELECT d.idDoctorant, p. cntPubli
FROM (SELECT idPersonnel, COUNT(idPublication) as cntPubli FROM PersonnelPublie GROUP BY idPersonnel) as

PRIGHT JOIN (SELECT idDoctorant FROM Doctorant) as d
ON d.idDoctorant=p.idPersonnel
```

Nous obtenons le résultat suivant :



A.1.6 R6

```
Le nombre de publications par annee de tout le laboratoire

(SELECT annee_publication,COUNT(*) as cntPubli FROM Publication GROUP BY annee_publication);
```

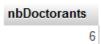
Nous obtenons le résultat suivant :

annee_publication	cntPubli
2018	5
2019	3

A.1.7 R7

```
Le nombre de doctorants du laboratoire

SELECT COUNT(*) as nbDoctorants FROM Doctorant;
```

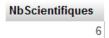


A.1.8 R8

```
Le nombre de scientifiques du laboratoire

SELECT COUNT(*) as NbScientifiques FROM Scientifique;
```

Nous obtenons le résultat suivant :



A.1.9 R9

```
--9. Le nombre d'enseignants chercheurs par établissement d'enseignement

(SELECT idEtablissement ,COUNT(*) as cntEnseignants FROM Enseignant_chercheur GROUP BY idEtablissement)
```

Nous obtenons le résultat suivant :

idEtablissement	cntEnseignants
51	2
52	1

A.1.10 R10

```
-- 10. Le nombre de publications par scientifique/doctorant

(SELECT idPersonnel, COUNT(idPublication) as cntPubli FROM PersonnelPublie GROUP BY idPersonnel)
```

Nous obtenons le résultat suivant :



A.1.11 R11

Nous obtenons le résultat suivant :



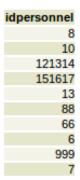
A.1.12 R12

```
--12. Les personnes qui n ont jamais participe aux journees portes ouvertes

(SELECT idPersonnel from Personnel)

EXCEPT

(SELECT idPersonnel from PersonnelParticipeJPO)
```



A.1.13 R13

Nous obtenons le résultat suivant :



A.1.14 R14

Nous obtenons le résultat suivant :



A.1.15 R15

Pour une annee donnee, on veut recuperer le nombre de publications, de conferences, et de $doctorants\ de\ chaque\ scientifique$ SELECT s.idScientifique, NbPubli, NbConf, NbDocts 3 FROM(SELECT id Scientifique FROM Scientifique) as s 4 LEFT JOIN (SELECT idPresident, COUNT(*) as NbConf FROM Conference WHERE EXTRACT(YEAR FROM date_debut) =2018 GROUP BY idPresident) as c ON s.idScientifique=c.idPresident 6 LEFT JOIN (SELECT idPersonnel, COUNT(*) as NbPubli FROM PersonnelPublic WHERE idPublication IN (SELECT idPublication FROM Publication WHERE annee publication=2018) GROUP BY idPersonnel) as p ON s.idScientifique=p.idPersonnel 9 10 (SELECT idScientifique, COUNT(*) as NbDocts FROM ScientifiqueEncadreDoctorant WHERE idDoctorant IN(SELECT idDoctorant FROM Doctorant WHERE EXTRACT(YEAR FROM date soutenance) = 2018) GROUP BY id Scientifique) as d ON s.idScientifique=d.idScientifique

idScientifique	NbPubli	NbConf	NbDocts
6	NULL	NULL	1
7	NULL	NULL	NULL
10	2	NULL	NULL
13	NULL	NULL	NULL
999	1	1	NULL
118518	NULL	NULL	NULL

A.1.16 R16

Nous obtenons le résultat suivant :

nom prenom Kine Miss

A.1.17 R17

```
--17. Afficher pour chaque scientifique, le nombre de ses publications, le nombre de ses projets et de ses doctorants.

SELECT s.idScientifique, NbPubli, NbProjet, NbDocts
FROM(SELECT idScientifique FROM Scientifique) as s
LEFT JOIN (SELECT idPersonnel, COUNT(*) as NbPubli FROM PersonnelPublie CROUP BY idPersonnel) as p
ON s.idScientifique=p.idPersonnel
LEFT JOIN (SELECT idScientifique, COUNT(*) as NbProjet FROM ScientifiqueParticipeProjet CROUP BY idScientifique) as pr

ON s.idScientifique=pr.idScientifique
LEFT JOIN (SELECT idScientifique, COUNT(*) as NbDocts FROM ScientifiqueEncadreDoctorant CROUP BY idScientifique) as d
ON s.idScientifique=d.idScientifique
```

Nous obtenons le résultat suivant :

idScientifique	NbPubli	NbProjet	NbDocts
6	NULL	NULL	3
7	1	NULL	1
10	3	1	2
13	NULL	NULL	NULL
999	1	NULL	NULL
118518	NULL	2	2

A.1.18 R18

```
idscientifique
-----
(0 rows)
```

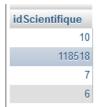
A.1.19 R19

Nous obtenons le résultat suivant :



A.1.20 R20

Nous obtenons le résultat suivant :



A.1.21 R21

```
--21. Le nombre de collaborateurs par pays

SELECT pn.Pays, COUNT(idAuteur)

FROM(SELECT Pays,Nom FROM Labo_externe) as pn

LEFT JOIN(SELECT idAuteur, NomLabo FROM Auteur) as a

ON pn.Nom=a.NomLabo

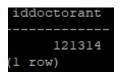
GROUP BY pn.Pays
```

pays	count
France	3
URSS	1
Estonie	1

A.1.22 R22

```
-22.
             Les doctorants qui ont un seul encadrant et qui ont toujours des publications qu avec leur
        encadrant
   SELECT idDoctorant FROM Doctorant
   WHERE idDoctorant IN (
   SELECT idPersonnel FROM PersonnelPublie
   WHERE idPublication IN (
   (SELECT idPublication FROM PersonnelPublie
   WHERE idPersonnel in
   (SELECT id Scientifique FROM Scientifique Encadre Doctorant
10
   CROUP BY id Scientifique)
11
   INTERSECT
13
14
   SELECT idPublication FROM PersonnelPublie
15
   WHERE idPersonnel in
16
   ( \textbf{SELECT} \  \, \text{idDoctorant} \  \, \textbf{FROM} \  \, \text{ScientifiqueEncadreDoctorant}
17
   GROUP BY id Doctorant
18
   HAVING COUNT(*)=1)))
```

Nous obtenons le résultat suivant :



A.1.23 R23

```
---23. Les doctorants qui ont plus de 3 encadrants

SELECT idDoctorant FROM ScientifiqueEncadreDoctorant GROUP BY idDoctorant HAVING COUNT(idScientifique)

>3
```

Nous obtenons le résultat suivant :



A.1.24 R24

Nous n'avons malheureusement pas pu tester cette requête, car le jeu de données n'était pas suffisant pour celle-ci.

A.1.25 R25

```
--25. Les doctorants qui ont au moins une publication chaque anne depuis leur recrutement

SELECT np.idPersonnel

FROM
(SELECT DISTINCT dp.idPersonnel, COUNT(pa.annee_publication) as ap
```

```
FROM(SELECT idPersonnel, idPublication FROM PersonnelPublie
WHERE idPersonnel IN (SELECT idDoctorant FROM Doctorant)) as dp

JOIN

(SELECT idPublication, annee_publication FROM Publication) as pa

ON dp.idPublication=pa.idPublication

GROUP BY dp.idPersonnel) as np

JOIN

(SELECT idPersonnel, EXTRACT(YEAR FROM CURRENT_DATE)-EXTRACT(YEAR FROM Date_recrutement) as Nbannees
FROM Personnel) as tp

ON tp.idPersonnel=np.idPersonnel

WHERE np.ap = tp.Nbannees
```



A.1.26 R26

```
-26.
           Les scientifiques qui recrutent au moins un doctorant par annee
1
3
   SELECT da. id Scientifique FROM
4
   (SELECT sd.idScientifique, COUNT(sd.DateDoc) as NbDocs, AVG(EXTRACT( YEAR FROM p.Date recrutement)) as
       DateSci
   FROM (SELECT idPersonnel, Date_recrutement FROM Personnel) as p
   JOIN (SELECT DISTINCT s.idScientifique, EXTRACT(YEAR FROM d.Date recrutement) as DateDoc
   FROM
   (SELECT idScientifique, idDoctorant FROM ScientifiqueEncadreDoctorant) as s
10
   JOIN
11
   (SELECT idPersonnel, Date recrutement FROM Personnel WHERE idPersonnel IN (SELECT idDoctorant FROM
       Doctorant)) as d
   ON s.idDoctorant=d.idPersonnel ) as sd
13
   ON sd.idScientifique=p.idPersonnel
   GROUP BY sd.idScientifique) as da
15
   WHERE (da.nbdocs>= (EXTRACT(YEAR FROM CURRENT DATE) - da.DateSci))
```

Nous obtenons le résultat suivant :

idscientifique 6

A.1.27 R27

```
-27. Les pays qui sont prents tous les projets

SELECT pc.Pays
FROM (SELECT DISTINCT COUNT(ppp. TitreProjet) as nbProj, pa.Pays
FROM
(SELECT DISTINCT Pays, idPartenaire FROM Partenaire) as pa
JOIN
(SELECT ** FROM PartenaireParticipeProjet) as ppp

ON ppp.idPartenaire=pa.idPartenaire
CROUP BY pa.Pays) as pc

WHERE pc.nbProj=(SELECT COUNT(*) FROM Projet recherche)
```

Pays France

A.1.28 R28

```
-28. Les scientifiques qui publient que dans des confrences de classe A

SELECT idScientifique FROM Scientifique WHERE idScientifique IN(
(SELECT idPersonnel FROM PersonnelPublie WHERE idPublication IN (SELECT idPublication FROM Publication WHERE nom_conference_journal IN (SELECT nom_conference_journal FROM TypeConf WHERE classe_conference='A')))

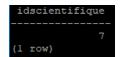
EXCEPT
(SELECT idPersonnel FROM PersonnelPublie WHERE idPublication IN (SELECT idPublication FROM Publication WHERE nom_conference_journal IN (SELECT nom_conference_journal FROM TypeConf WHERE classe_conference!='A'))))
```

Nous obtenons le résultat suivant :

idscientifique 999

A.1.29 R29

Nous obtenons le résultat suivant :



A.1.30 R30

annee	nbconf
2019	1
2018	2
2001	1
2000	1

A.1.31 R31

```
--31. L tablissement denseignement ayant le plus grand nombre denseignant chercheur

SELECT ident.idetablissement FROM

(SELECT idetablissement COUNT(*) as entenseignants
FROM Enseignant_chercheur
GROUP BY idetablissement) as ident

WHERE ident.entEnseignants=(SELECT max(nc.entEnseignants))
FROM (
SELECT idetablissement COUNT(*) as entenseignants
FROM Enseignant_chercheur
GROUP BY idetablissement) as nc)
```

