### ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

ΕΞΑΜΗΝΙΑΙΑ ΕΡΓΑΣΙΑ

Ακαδημαϊκό έτος 2021-2022

ΟΜΑΔΑ 60

Νικόλαος Καραφύλλης Α.Μ.:03119890

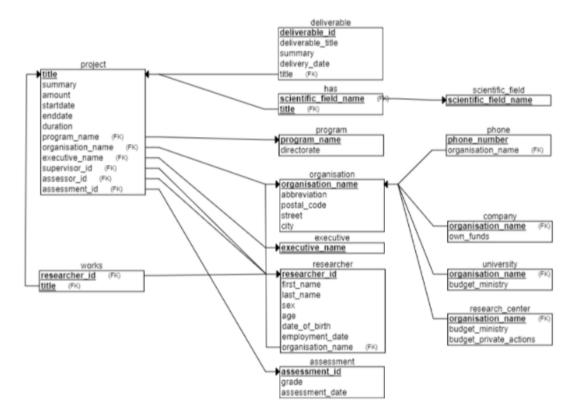
Αναστάσιος Σαφαρής: | Α.Μ.:03119896

Ιγνάτιος Σιμινής: | Α.Μ.:03119821

# Σύστημα αποθήκευσης, διαχείρισης και ανάλυσης πληροφοριών ΕΛ.Ι.Δ.Ε.Κ

#### 2.1

Παρατίθεται το σχεσιακό διάγραμμα που χρησιμοποιήθηκε και αντιστοιχεί στο ER διάγραμμα που μας δόθηκε.



### **INDEXES**

Ακόμα, έχουμε ορίσει 5 κατάλληλα ευρετήρια (indexes) τα οποία έχουν οριστεί με σκοπό την ελαχιστοποίηση του χρόνου που απαιτείται για την πραγματοποίηση των queries. Η

συγκεκριμένη επιλογή έγινε, καθώς τα συγκεκριμένα attributes καλούνται με μεγάλη συχνότητα στα ερωτήματα.

```
1    CREATE INDEX idx_Startdate ON project(startdate);
2    CREATE INDEX idx_Enddate ON project(enddate);
3    CREATE INDEX idx_resfirstname ON researcher(first_name);
4    CREATE INDEX idx_reslastname ON researcher(last_name);
5    CREATE INDEX idx_dateofbirth ON researcher(date_of_birth);
```

2.2

Query που χρησιμοποιήθηκε για την δημιουργία της βάσης δεδομένων

```
DROP DATABASE IF EXISTS `HFRIManagement`;
CREATE DATABASE `HFRIManagement`;
ALTER DATABASE HFRIManagement
CHARACTER SET utf8
COLLATE utf8_bin;
```

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα project

```
CREATE TABLE project
   title varchar (50) PRIMARY KEY not null,
   summary varchar (1000),
   amount int CHECK (amount>=100000 AND amount<=1000000),
   startdate date,
   enddate date,
   program_name varchar(50),
   organisation_name varchar (50),
   executive_name varchar(50),
   supervisor_id int,
    assessment_id int,
   CONSTRAINT duration CHECK ((timestampdiff(year,startdate, enddate))>=1 and (timestampdiff(year,enddate, startdate))<=4),
   FOREIGN KEY (program_name) REFERENCES program(program_name) ON DELETE SET NULL ON UPDATE CASCADE,
    FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE SET NULL ON UPDATE CASCADE,
    FOREIGN KEY (executive_name) REFERENCES executive(executive_name) ON DELETE SET NULL ON UPDATE CASCADE,
    FOREIGN KEY (supervisor_id) REFERENCES researcher(researcher_id) ON DELETE SET NULL ON UPDATE CASCADE,
    FOREIGN KEY (assessor_id) REFERENCES researcher(researcher_id) ON DELETE SET NULL ON UPDATE CASCADE,
    FOREIGN KEY (assessment_id) REFERENCES assessment(assessment_id) ON DELETE SET NULL ON UPDATE CASCADE
```

Για primary key επιλέχθηκε το attribute title, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε έργου. Το program\_name αποτελεί foreign key για τον πίνακα program και αντιστοιχίζει τη σχέση πολλά (project) σε ένα (program). Το organisation\_name αποτελεί foreign key για τον πίνακα organisation και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (project) σε ένα (organisation). Το supervisor\_id και το assessor\_id αποτελούν foreign keys για τον πίνακα researcher και αντιστοιχίζουν τους δύο πίνακες με σχέση πολλά (project) σε ένα (researcher). Το executive\_name αποτελεί foreign key για τον πίνακα executive και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (project) σε ένα (executive). Το assessment\_id αποτελεί foreign key για τον πίνακα assessment και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (project) σε ένα (assessment). Τέλος, έχουμε βάλει ένα CHECK constraint που να ελέγχει ότι η διάρκεια ενός project είναι από 1 έως 4 χρόνια.

#### Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα deliverable

```
DROP TABLE IF EXISTS deliverable;

CREATE TABLE deliverable

(

    deliverable_id int auto_increment PRIMARY KEY NOT NULL,
    deliverable_title varchar (50) not null,
    summary varchar (500),
    delivery_date date,
    title varchar(50) NOT NULL,
    FOREIGN KEY (title) REFERENCES project(title) ON DELETE CASCADE ON UPDATE CASCADE
);
```

Για primary key επιλέχθηκε το attribute deliverable\_id, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε παραδοτέου. Το title αποτελεί foreign key για τον πίνακα project και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (deliverable) σε ένα (project).

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα program

```
DROP TABLE IF EXISTS program;

CREATE TABLE program

(
    program_name varchar (50) PRIMARY KEY not null,
    directorate varchar (100)
);
```

Για primary key επιλέχθηκε το attribute program\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό του κάθε προγράμματος.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα organisation

```
DROP TABLE IF EXISTS organisation;

CREATE TABLE organisation

(
    organisation_name varchar (50) PRIMARY KEY not null,
    abbreviation varchar (10),
    postal_code varchar (50),
    street varchar (50),
    city varchar (50)

);
```

Για primary key επιλέχθηκε το attribute organisation\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό του κάθε οργανισμού.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα executive

```
DROP TABLE IF EXISTS executive;

CREATE TABLE executive

(
    executive_name varchar (50) PRIMARY KEY not null
);
```

Για primary key επιλέχθηκε το attribute executive\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε στελέχους. Άλλωστε, είναι το μοναδικό attribute της συγκεκριμένης οντότητας.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα researcher

```
DROP TABLE IF EXISTS researcher;

CREATE TABLE researcher

(

researcher_id int auto_increment PRIMARY KEY NOT NULL,

first_name varchar (50) not null,

last_name varchar (50) not null,

sex varchar (12) not null,

date_of_birth date,

/*age int GENERATED ALWAYS AS (TIMESTAMPDIFF(YEAR, date_of_birth, current_date())),*/ /*mono sta queries ta derived*/

employment_date date,

organisation_name varchar (50) ,

FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE SET NULL ON UPDATE CASCADE

);
```

Για primary key επιλέχθηκε το attribute researcher\_id, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε ερευνητή. Το organisation\_name αποτελεί foreign key για τον πίνακα organisation και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (researcher) σε ένα (organisation).

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα assessment

```
DROP TABLE IF EXISTS assessment;

CREATE TABLE assessment

(
    assessment_id int auto_increment PRIMARY KEY NOT NULL,
    grade int not null,
    assessment_date date
);
```

Για primary key επιλέχθηκε το attribute assessment\_id, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε αξιολόγησης.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα scientific field

```
DROP TABLE IF EXISTS scientific_field;

CREATE TABLE scientific_field

(
    scientific_field_name varchar (50) PRIMARY KEY not null
);
```

Για primary key επιλέχθηκε το attribute scientific\_field\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε επιστημονικού πεδίου. Άλλωστε, είναι το μοναδικό attribute της συγκεκριμένης οντότητας.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα phone

```
DROP TABLE IF EXISTS phone;

CREATE TABLE phone

(
    phone_number varchar (50) PRIMARY KEY NOT NULL,
    organisation_name varchar (50) NOT NULL,
    FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE CASCADE ON UPDATE CASCADE
);
```

Για primary key επιλέχθηκε το attribute phone\_number, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε τηλεφώνου. Το organisation\_name αποτελεί foreign key στον πίνακα organisation και αντιστοιχίζει τους δύο πίνακες με σχέση πολλά (phone) σε ένα (organisation).

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα company

```
DROP TABLE IF EXISTS company;
CREATE TABLE company

(
    organisation_name varchar (50) PRIMARY KEY NOT NULL,
    own_funds int not null,
    FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE CASCADE ON UPDATE CASCADE
);
```

Για primary key επιλέχθηκε το attribute organisation\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε εταιρείας.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα university

```
DROP TABLE IF EXISTS university;

CREATE TABLE university

(
    organisation_name varchar (50) PRIMARY KEY NOT NULL,
    budget_ministry int not null,
    FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE CASCADE ON UPDATE CASCADE
);
```

Για primary key επιλέχθηκε το attribute organisation\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε πανεπιστημίου.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα research center

```
DROP TABLE IF EXISTS research_center;

CREATE TABLE research_center

(
    organisation_name varchar (50) PRIMARY KEY NOT NULL,
    budget_ministry int not null,
    budget_private_actions int not null,
    FOREIGN KEY (organisation_name) REFERENCES organisation(organisation_name) ON DELETE CASCADE ON UPDATE CASCADE
);
```

Για primary key επιλέχθηκε το attribute organisation\_name, διότι αυτό είναι μοναδικό και χαρακτηριστικό κάθε ερευνητικού κέντρου.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα works

Ο πίνακας works είναι μία σχέση (πολλά προς πολλά) που ενώνει τις οντότητες project και researcher. Για αυτό, εκτός από το primary key (researcher\_id, title) που είναι composite key και χαρακτηρίζει μοναδικά κάθε εγγραφή του, περιέχει και το researcher\_id ως foreign key για το researcher και το title ως foreign key για το project.

Query που χρησιμοποιήθηκε για την δημιουργία του πίνακα has

```
DROP TABLE IF EXISTS has;

CREATE TABLE has

(

    title varchar(50) NOT NULL,
    scientific_field_name varchar(50) NOT NULL,
    PRIMARY KEY (title, scientific_field_name),
    FOREIGN KEY (title) REFERENCES project(title) ON DELETE CASCADE ON UPDATE CASCADE,
    FOREIGN KEY (scientific_field_name) REFERENCES scientific_field(scientific_field_name)
    ON DELETE CASCADE ON UPDATE CASCADE
);
```

Ο πίνακας has είναι μία σχέση (πολλά προς πολλά) που ενώνει τις οντότητες project και scientific\_field. Για αυτό, εκτός από το primary key (scientific\_field\_name, title) που είναι composite key και χαρακτηρίζει μοναδικά κάθε εγγραφή του, περιέχει και το scientific\_field\_name ως foreign key για το researcher και το title ως foreign key για το project.

#### TRIGGERS

Προκειμένου να καλύψουμε τον περιορισμό της εκφώνησης ότι ο αξιολογητής κάθε έργου δεν μπορεί να ανήκει στον οργανισμό που διαχειρίζεται το συγκεκριμένο έργο, έχουμε δημιουργήσει δύο triggers (ένα before insert και ένα before update) τα οποία ελέγχουν τη συγκεκριμένη συνθήκη και εμφανίζουν κατάλληλο μήνυμα σφάλματος.

```
DELIMITER $$

DROP TRIGGER IF EXISTS before_assessor_insert$$

CREATE TRIGGER before_assessor_insert

before insert ON project FOR EACH ROW

BEGIN

IF new.assessor_id in (select a.researcher_id from researcher a where new.organisation_name = a.organisation_name)

THEN SIGNAL SQLSTATE '50001' SET MESSAGE_TEXT = 'Assessor cannot assess a project belonging in his organisation';

END IF;

END $$

DELIMITER $$

DROP TRIGGER IF EXISTS before_assessor_update$$

CREATE TRIGGER before_assessor_update

before update ON project FOR EACH ROW

BEGIN

IF new.assessor_id in (select a.researcher_id from researcher a where new.organisation_name = a.organisation_name)

THEN SIGNAL SQLSTATE '50004' SET MESSAGE_TEXT = 'New assessor cannotassess a project belonging in his organisation';

END IF;

END IF;

END S$
```

#### **USERS**

Έχουμε δημιουργήσει 3 users με διαβαθμισμένα δικαιώματα στη βάση.

```
CREATE USER superuser@localhost;
GRANT ALL PRIVILEGES ON hfrimanagement.* TO superuser;

CREATE USER data_entry@localhost;
GRANT SELECT, INSERT, DELETE, UPDATE ON hfrimanagement.* TO data_entry;

CREATE USER viewer@localhost;
GRANT SELECT ON hfrimanagement.* TO viewer;
```

## ΡΥΤΗΟΝ ΚΩΔΙΚΑΣ-ΕΙΣΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ

Ο κώδικας (insert\_data.py) που χρησιμοποιήθηκε για την εισαγωγή των dummy δεδομένων στην βάση από τα αρχεία .csv είναι ο ακόλουθος:

```
import mysql.connector
     import pandas as pd
     import numpy as np
    programs = pd.read_csv("./Data/programs.csv")
    scientific fields= pd.read csv("./Data/scientific fields.csv")
    executives = pd.read_csv("./Data/executives.csv")
    assessments = pd.read_csv("./Data/assessments.csv")
    organisations = pd.read_csv("./Data/organisations.csv")
    organisation_types = pd.read_csv("./Data/organisation_type.csv")
    phone_numbers = pd.read_csv("./Data/phone_numbers.csv")
    researchers = pd.read_csv("./Data/researchers.csv")
13
    projects = pd.read_csv("./Data/projects.csv")
    has = pd.read_csv("./Data/has.csv")
    deliverables = pd.read csv("./Data/deliverable.csv")
16
    works = pd.read_csv("./Data/works.csv")
    mydb = mysql.connector.connect(
19
        host = "localhost",
        user = "root",
         passwd = '',
         database = "HFRIManagement"
```

```
mycursor.execute(sqlFormula)
    mydb.commit()
for i in range(len(executives)):
    executive_name = executives['executive_name'][i].replace("","").replace("","")
    sqlFormula = """INSERT INTO executive (executive_name)
                    VALUES ('{}')""".format(executive_name)
    mycursor.execute(sqlFormula)
   mydb.commit()
for i in range(len(assessments)):
    grade = assessments['grade'][i]
    assessment_date = assessments['assessment_date'][i].replace("'","").replace('"',""
    sqlFormula = """INSERT INTO assessment (grade,assessment_date)
                    VALUES ('{}','{}')""".format(grade,assessment_date)
    mycursor.execute(sqlFormula)
    mydb.commit()
for i in range(len(organisations)):
```

```
organisation_name = organisations['organisation_name'][i].replace("","").replace("","")

abbreviation = organisations['abbreviation'][i].replace("","").replace("","")

postal_code = organisations['postal_code'][i]

street= organisations['street'][i].replace("","").replace("","")

city= organisations['city'][i].replace("","").replace("","")

sqlFormula = ""INSERT INTO organisation (organisation_name,abbreviation,postal_code,street,city)

VALUES ('{}','{}','{}','{}','{}','",'','',''',''','''''''.format(organisation_name,abbreviation,postal_code,street,city)

mycursor.execute(sqlFormula)

mydb.commit()

for i in range(len(organisation_types)):

own_funds = organisation_types['budget_ministry'][i]

budget_ministry= organisation_types['budget_ministry'][i]

budget_ministry2= organisation_types['budget_ministry2'][i]

organisation_name-organisation_types['budget_ministry2'][i]

sqlFormula1 = """INSERT INTO company (own_funds,organisation_name)

VALUES ('{}','{}')'"".format(own_funds,organisation_name)

VALUES ('{}','{}')'"".format(budget_ministry,organisation_name)

sqlFormula2 = """INSERT INTO research_center (budget_private_actions,budget_ministry,organisation_name)

sqlFormula3 = """INSERT INTO research_center (budget_private_actions,budget_ministry,organisation_name)
```

```
VALUES ('{}','{}','{}')""".format(budget_private_actions,budget_ministry2,organisation_name
      mycursor.execute(sqlFormula1)
      mycursor.execute(sqlFormula2)
      mycursor.execute(sqlFormula3)
      mydb.commit()
for i in range(len(phone_numbers)):
     phone_number = phone_numbers['phone_number'][i].replace("","").replace("","")
organisation_name = phone_numbers['organisation_name'][i]
     mycursor.execute(sqlFormula)
     mydb.commit()
for i in range(len(researchers)):
     first_name = researchers['first_name'][i].replace("","").replace("","")
last_name = researchers['last_name'][i].replace("","").replace(""',"")
sex = researchers['sex'][i].replace("","").replace(""',"")
date_of_birth = researchers['date_of_birth'][i]
      employment_date = researchers['employment_date'][i]
     organisation name = researchers['organisation name'][i]
     sqlFormula = """INSERT INTO researcher (first_name,last_name,sex,date_of_birth,employment_date,organisation_name)
VALUES ('{}','{}','{}','{}','{}')""".format(first_name,last_name,sex,date_of_birth,
                           employment_date,organisation_name)
     mycursor.execute(sqlFormula)
     mydb.commit()
for i in range(len(projects)):
    title = projects['title'][i].replace("'","").replace('"',"")
    summary = projects['summary'][i].replace("'","").replace('"',"")
    amount = projects['amount'][i]
    startdate = projects['startdate'][i].replace("'","").replace('"',"")
    enddate = projects['enddate'][i].replace("'","").replace('"',"")
    #duration = timestampdiff(year,enddate, startdate)
    program_name = projects['program_name'][i]
    organisation_name = projects['organisation_name'][i]
     organisation_name = projects['organisation_name'][i]
     executive_name = projects['executive_name'][i]
     supervisor_id = projects['supervisor_id'][i]
     assessor_id = projects['assessor_id'][i]
     assessment_id = projects['assessment_id'][i]
      sqlFormula = """INSERT INTO project (title, summary, amount, startdate, enddate, program_name, organisation_name,
                           program_name,organisation_name,executive_name,supervisor_id,assessor_id,assessment_id)
      mycursor.execute(sqlFormula)
      mydb.commit()
for i in range(len(has)):
      scientific_field_name = has['scientific_field_name'][i]
      title = has['title'][i]
      mycursor.execute(sqlFormula)
     mydb.commit()
for i in range(len(deliverables)):
     deliverable_title = deliverables['deliverable_title'][i].replace("","").replace("","")
summary= deliverables['summary'][i].replace("","").replace("","")
delivery_date = deliverables['delivery_date'][i].replace("","").replace('"',"")
```

```
delivery_date = deliverables['delivery_date'][i].replace(""","").replace('"","")

title = deliverables['title'][i]

sqlFormula = """INSERT INTO deliverable (deliverable_title, summary, delivery_date, title)

sqlFormula = """INSERT INTO deliverable (deliverable_title, summary, delivery_date, title)

vALUES ('{}', '{}', '{}', ''{}', '''"".format(deliverable_title, summary, delivery_date, title)

mycursor.execute(sqlFormula)
mydb.commit()

for i in range(len(works)):
    researcher_id = works['researcher_id'][i]

title= works['title'][i]
```

```
sqlFormula = """INSERT INTO works (researcher_id,title)

VALUES ('{}','{}')""".format(researcher_id,title)

mycursor.execute(sqlFormula)
mydb.commit()

mydb.commit()

mydb.close()

mydb.close()

print("All Done! Bye, for now.")
```

2.3

# Configuration για την εγκατάσταση της εφαρμογής

- 1. Απαιτείται εγκατάσταση του git.
- 2. Έπειτα χρειάζεται να κάνουμε git clone το repository.
- 3. Απαιτείται η εγκατάσταση των Python 3, XAMPP Control Panel.
- 4. Από το XAMPP πατάμε start για το module MySQL.
- 5. Εκτέλεση της εντολής "pip install -r requirements.txt" ώστε να εγκατασταθούν τα υπόλοιπα απαιτούμενα.
- Στο shell του ΧΑΜΡΡ τρέχουμε τον κώδικα CREATE\_tables.sql, CREATE\_views.sql, CREATE\_indexes.sql.
- 7. Για την εισαγωγή των dummy δεδομένων στη βάση εκτελούμε την εντολή "python insert data.py".
- 8. Εκτελούμε την εντολή "python run.py" ώστε να ξεκινήσει η εφαρμογή.
- 9. Ανοίγουμε σε ένα browser τη διεύθυνση localhost:3000 ώστε να δούμε την εφαρμογή.

2.4

Σύνδεσμος για το git repo της εφαρμογής:

https://github.com/nickarafyllis/Elidek-database