DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

Databases Project – Spring 2017

Prof. Anastasia Ailamaki

Team No: 15

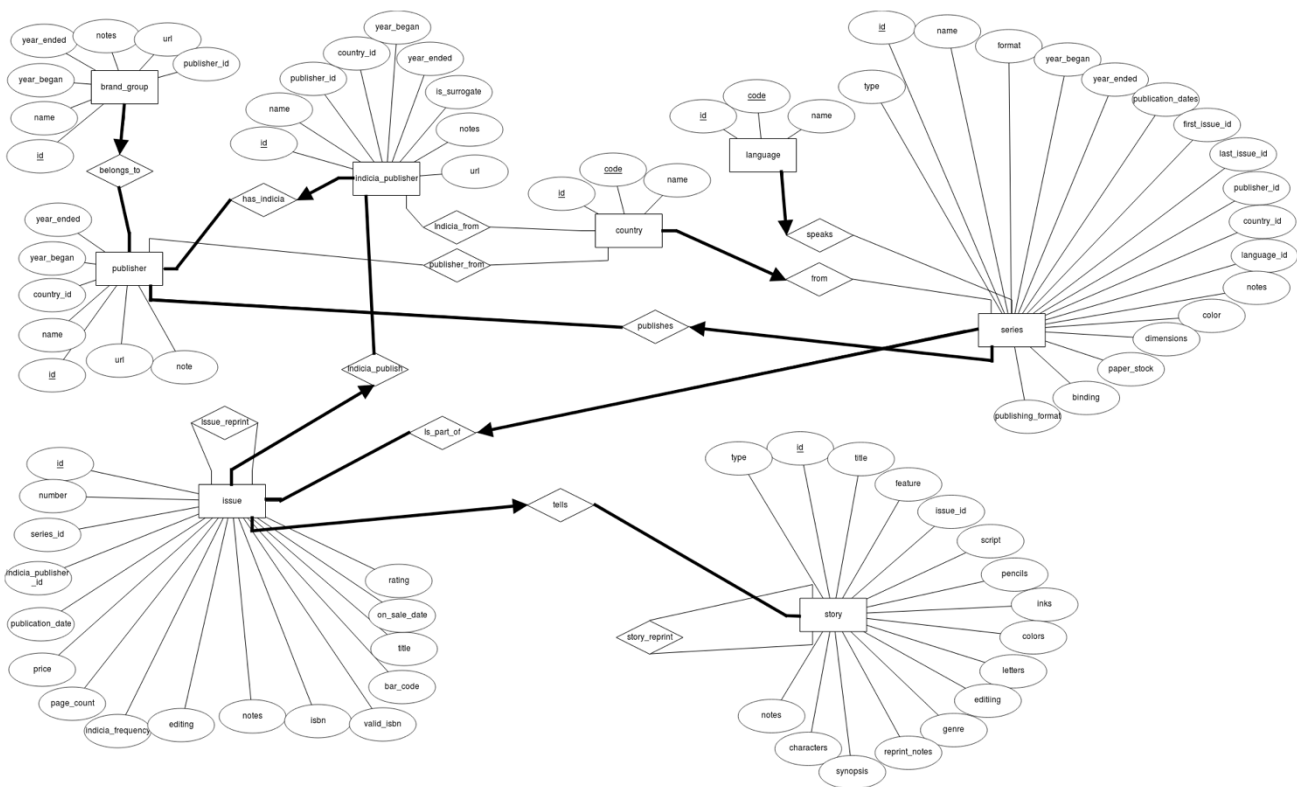Names: Colin Branca, Jules Courtois, Yoan Martin

Contents

Deliverable 1

Assumptions

When we create the tables, we assume that each CHAR has a max size. No longer elements will be added later.

Entity Relationship Schema

Schema

Description

There are two main parts in the schema: Books with Story, Issue and Series and Company with Publisher, Indicia Publisher and Brand Group. Let's first describe each part before explaining the connections between them.

Books: This part is describing a book in the more general sense. The physical book is an issue. It tells a story and it is part of a series. For example, Harry Potter is a famous series. Harry Potter and the Philosopher's Stone is a story from this series. Finally, the book with ISBN X is a physical book telling this story. Since these three components are highly connected. There are a lot of constraints between them.

1) An issue tells at least a story.
2) A story must be told by exactly one issue.

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

3) Each series has at least one issue.
4) An issue is part of exactly one series.

Company: This part is describing the company which publish some books. The main company is the publisher. It has smallest companies, the indicia publishers and it holds some brand group. For example, Marvel is the general company. It holds a brand group called Thor and the physical books are published by an indicia publisher, Thor Entertaining Group, which is part of Marvel. As for the book part, the company part has also a lot of constraints.
1) Each publisher has at least one brand group.
2) Each brand group belongs to exactly one publisher.
3) Each publisher has at least one indicia publisher.
4) Each indicia publisher is part of exactly one publisher.

Then, there are some connections between these two parts. The publisher publishes a series. It means that the publisher creates a series but the issues are printed by the indicia publisher. Here are the constraints related to these connections.
1) A publisher publishes at least one series.
2) A series is published by exactly one publisher.
3) An indicia publisher publishes at least one issue.
4) An issue is published by exactly one indicia publisher.

Finally, there are some smallest relationships. Firstly, the publisher, the indicia publisher and the series come from a country and the series have a language. Secondly, there is a reprint relationship between two stories or two issues. We consider only two constraints.
1) A publisher, indicia publisher or series comes from exactly one country.
2) A series has exactly one language.

N.B. We do not create a story_type and series_publication_type entity since these data contain only one attribute if we exclude the id. To simplify the schema, we add an attribute type containing the name of the type in the Story/Series entity.

Relational Schema

ER schema to Relational schema

To create the table, we simply took each entity or each relationship and we translate them into a table. We took care to respect the constraints using NOT NULL or PRIMARY KEY to state the "at least" constraint and we permit each "at most" relation to appear in only one column in the connected table.

DDL

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

```sql
CREATE TABLE StoryReprint (
  id INT,
  origin_id INT NOT NULL,
  target_id INT NOT NULL,
  PRIMARY KEY (id)
);

CREATE TABLE IssueReprint (
  id INT,
  origin_issue_id INT NOT NULL,
  target_issue_id INT NOT NULL,
  PRIMARY KEY (id)
);

CREATE TABLE Country (
  id INT,
  code CHAR(4) NOT NULL,
  name CHAR(36) NOT NULL,
  PRIMARY KEY (id),
  UNIQUE (code),
  UNIQUE (name)
);
```

```sql
CREATE TABLE Story (
  id INT,
  type VARCHAR(50) NOT NULL,
  title VARCHAR(50) NOT NULL,
  feature VARCHAR(50),
  issue_id INT NOT NULL,
  script VARCHAR(50),
  pencils VARCHAR(50),
  inks VARCHAR(50),
  color VARCHAR(50),
  letters VARCHAR(50),
  editing VARCHAR(50),
  genre VARCHAR(50),
  characters VARCHAR(50),
  synopsis VARCHAR(150),
  reprint_notes VARCHAR(50),
  notes VARCHAR(50),
  PRIMARY KEY (id),
  FOREIGN KEY (issue_id) REFERENCES issue (id)
);
```

```sql
CREATE TABLE Issue (
  id INT,
  issue_number INT,
  series_id INT,
  indicia_publisher-id INT,
  publication_date CHAR(15),
  price CHAR(10),
  page_count INT,
  indicia_frequency CHAR(10),
  editing VARCHAR(20),
  notes VARCHAR(1000),
  isbn VARCHAR(20),
  valid_isbn VARCHAR(20),
  barcode VARCHAR(10),
  title VARCHAR(30),
  on_sale_date INT,
  rating INT,
  PRIMARY KEY (id),
  FOREIGN KEY (series_id) REFERENCES Series (id),
  FOREIGN KEY (indicia_publisher) REFERENCES IndiciaPublisher(id)
);
```

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

```sql
CREATE TABLE BrandGroup (
  id INT,
  name VARCHAR2(128) NOT NULL,
  year_began SMALLINT,
  year_ended SMALLINT,
  notes VARCHAR2(1023),
  url VARCHAR(128),
  publisher_id INT NOT NULL,
  PRIMARY KEY (id),
  UNIQUE (name)
);

CREATE TABLE Publisher (
  id INT,
  name VARCHAR2(128) NOT NULL,
  year_began SMALLINT,
  year_ended SMALLINT,
  notes VARCHAR2(3500),
  url VARCHAR(128),
  country_id INT,
  PRIMARY KEY (id),
  UNIQUE (name)
);

CREATE TABLE IndiciaPublisher (
  id INT,
  name VARCHAR2(128) NOT NULL,
  year_began SMALLINT,
  year_ended SMALLINT,
  notes VARCHAR2(3500),
  url VARCHAR(128),
  country_id INT,
  publisher_id INT,
  is_surrogate NUMBER(3) NOT NULL,
  PRIMARY KEY (id),
  UNIQUE (name)
);

CREATE TABLE Language (
  id INT,
  code CHAR(3) NOT NULL,
  name CHAR(36) NOT NULL,
  PRIMARY KEY (id),
  UNIQUE (code),
  UNIQUE (name)
);

CREATE TABLE Series (
  id INT,
  name VARCHAR2(1000) NOT NULL,
  type VARCHAR(16),
  format VARCHAR(30),
  publication_dates VARCHAR(200),
  year_began SMALLINT,
  year_ended SMALLINT,
  first_issue_id INT,
  last_issue_id INT,
  publisher_id INT,
  country_id INT,
  language_id INT,
  notes VARCHAR2(4000),
  color VARCHAR(200),
  dimensions VARCHAR(200),
  paper_stock VARCHAR(200),
  binding VARCHAR(200),
  publishing_format VARCHAR(200),
  PRIMARY KEY (id),
  UNIQUE (name),
  FOREIGN KEY (first_issue_id) REFERENCES Issue (id),
  FOREIGN KEY (last_issue_id) REFERENCES Issue (id),
  FOREIGN KEY (publisher_id) REFERENCES Publisher (id),
  FOREIGN KEY (country_id) REFERENCES Country (id),
  FOREIGN KEY (language_id) REFERENCES Language (id)
);
```

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

General Comments

For this deliverable, we design the schema together. Then we split the work individually. Colin Branca and Jules Courtois wrote the SQL commands to create the tables and Yoan Martin wrote this pdf document.