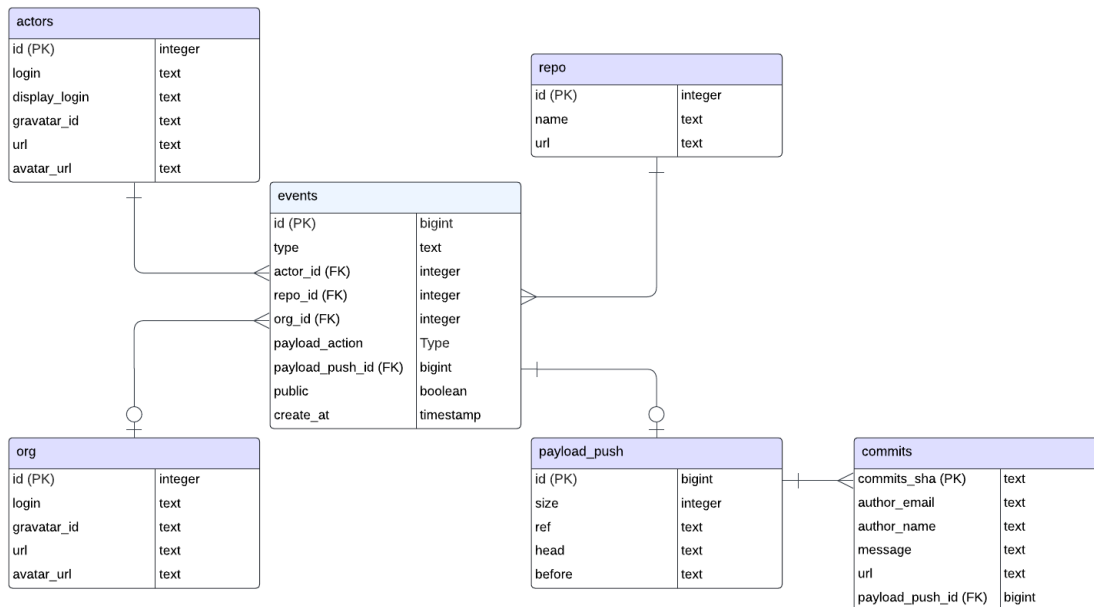


Project:	DS525 - Data Modeling I
Author:	Suthavee Piyavat (65199160203)

Data Model – ER Diagram



Project Instruction:

1. `$ cd 01-data-modeling-i/` เข้าไปที่ Folder 01-data-modeling-i
2. `$ docker-compose up` เชื่อมต่อ Database
3. เปิด Port 8080 เข้า Database ด้วยรหัสที่อยู่ในไฟล์ `docker-compose.yml`
4. `$ python create_tables.py` สร้าง Table บน Database

01-data-modeling-i > create_tables.py > ...

```

16
17 table_create_actors = """
18     CREATE TABLE IF NOT EXISTS actors (
19         id int,
20         login text,
21         display_login text,
22         gravatar_id text,
23         url text,
24         avatar_url text,
25         PRIMARY KEY(id)
26     )
27 """
28
29 table_create_repo = """
30     CREATE TABLE IF NOT EXISTS repo (
31         id int,
32         name text,
33         url text,
34         PRIMARY KEY(id)
35     )
36 """
  
```

Tables and views

Search data in tables (6)

 Search

<input type="checkbox"/>	Table	Engine	Collation	Data Length?	Index Length?	Data Free	Auto Increment	Rows?	Comment?
<input type="checkbox"/>	actors	table		0	16,384	?	?	0	
<input type="checkbox"/>	commits	table		0	16,384	?	?	0	
<input type="checkbox"/>	events	table		0	16,384	?	?	0	
<input type="checkbox"/>	org	table		0	16,384	?	?	0	
<input type="checkbox"/>	payload_push	table		0	16,384	?	?	0	
<input type="checkbox"/>	repo	table		0	16,384	?	?	0	
	6 in total		en_US.utf8	0	98,304	0			

5. \$ python etl.py โหลดข้อมูลจาก JSON File เข้า Table

01-data-modeling-i > etl.py > process

```
64     for datafile in all_files:
65         with open(datafile, "r") as f:
66             data = json.loads(f.read())
67             for each in data:
68
69                 insert_actors_val = (
70                     each["actor"]["id"],
71                     each["actor"]["login"],
72                     each["actor"]["display_login"],
73                     each["actor"]["gravatar_id"],
74                     each["actor"]["url"],
75                     each["actor"]["avatar_url"]
76                 )
77
78                 insert_repo_val = (
79                     each["repo"]["id"],
80                     each["repo"]["name"],
81                     each["repo"]["url"]
82                 )
83
84                 cur.execute(actors_table_insert, insert_actors_val)
85                 cur.execute(repo_table_insert, insert_repo_val)
```

Select: events

Select data Show structure Alter table New item

Select Search Sort Limit 50 Text length 100 Action Select

SELECT * FROM "events" LIMIT 50 (0.001 s) Edit

<input type="checkbox"/> Modify	id	type	actor_id	repo_id	org_id	payload_action	payload_push_id	public	create_at
<input type="checkbox"/> edit	23487929637	IssueCommentEvent	1696078	75340147	24305026	created	NULL	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929676	PushEvent	66924041	525526680	NULL	pushed	10754961780	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23488014712	PushEvent	106313266	521245400	NULL	pushed	10755001925	t	2022-08-17 15:55:05
<input type="checkbox"/> edit	23487929674	PushEvent	74971347	512700417	106303238	pushed	10754961752	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929661	PushEvent	91143053	525851446	NULL	pushed	10754961791	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929682	PushEvent	60825784	505029866	NULL	pushed	10754961801	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929673	PushEvent	102357035	525845306	NULL	pushed	10754961758	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929588	PushEvent	46447321	274724694	66193563	pushed	10754961749	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929636	CreateEvent	41898282	523780252	NULL	NULL	NULL	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929580	IssuesEvent	21972349	524701461	NULL	opened	NULL	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929591	PushEvent	17869732	164541747	NULL	pushed	10754961756	t	2022-08-17 15:51:05
<input type="checkbox"/> edit	23487929533	PushEvent	37043957	524655561	NULL	pushed	10754961711	t	2022-08-17 15:51:05