**Database final project**

Nikita Parfenov

[nikyparfenov@gmail.com](mailto:nikyparfenov@gmail.com)

**Goal of the project**

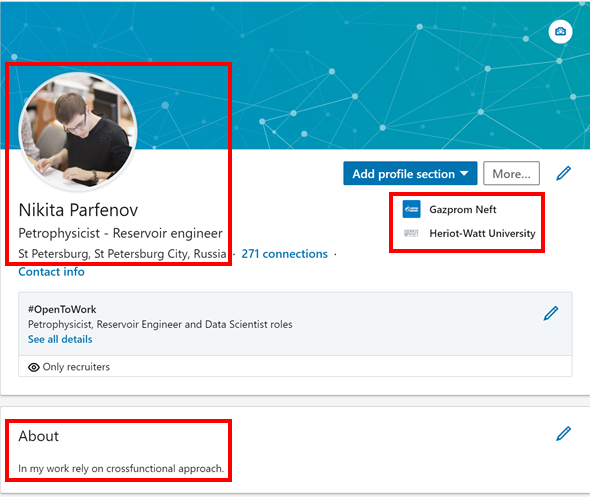
The purpose of the project is to describe simplified SQL database structure of popular social network LinkedIn ([https://www.linkedin.com](https://www.linkedin.com/)) and generate probable SQL queries that could be useful for the network data analysis.

The logic of database structure was manually separated into three block for analysis (user’s profile information, user’s social interaction, groups/companies profiles) and described below.

The project has been performed using MySQL database version 8.0.21 for Linux (Ubuntu).

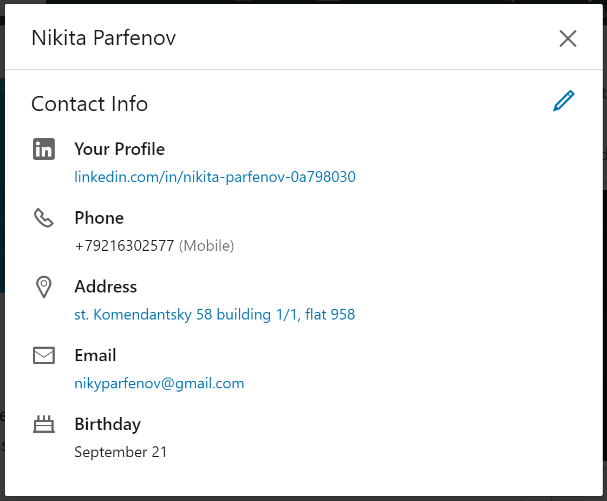
**Description and analysis of database structure**

**1. User’s profile information**

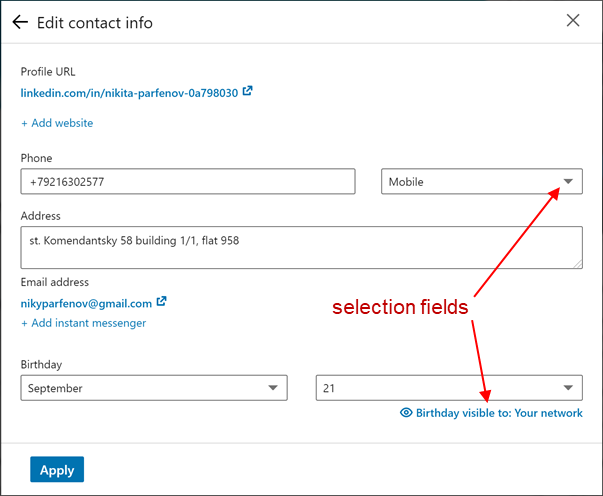


The general profile page contains user’s photo, first and last names, headline of user’s profession, city, country, information about current work place, educational institution and some block “About” with additional information.

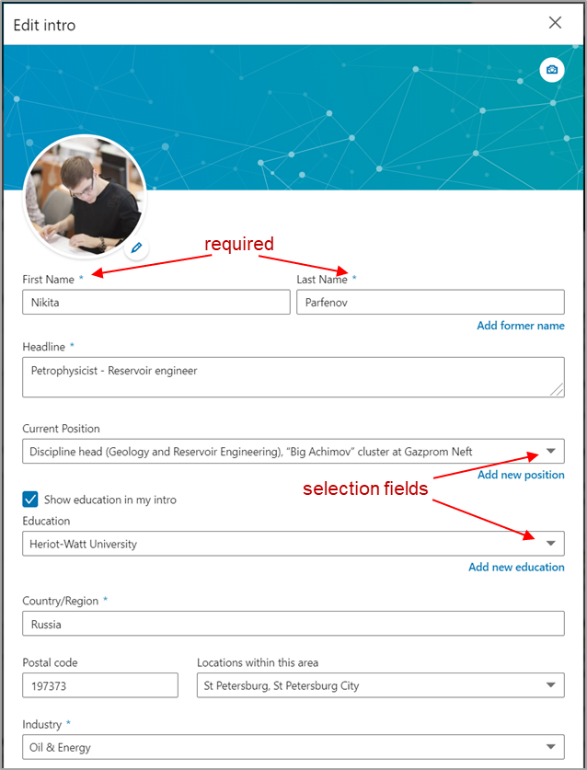
*\*To simplify the database structure job preferences page (#OpenToWork filed) has been omitted.*



Opening contact info, the information about url-adress of profile page, phone number, address, email and birthday appears.

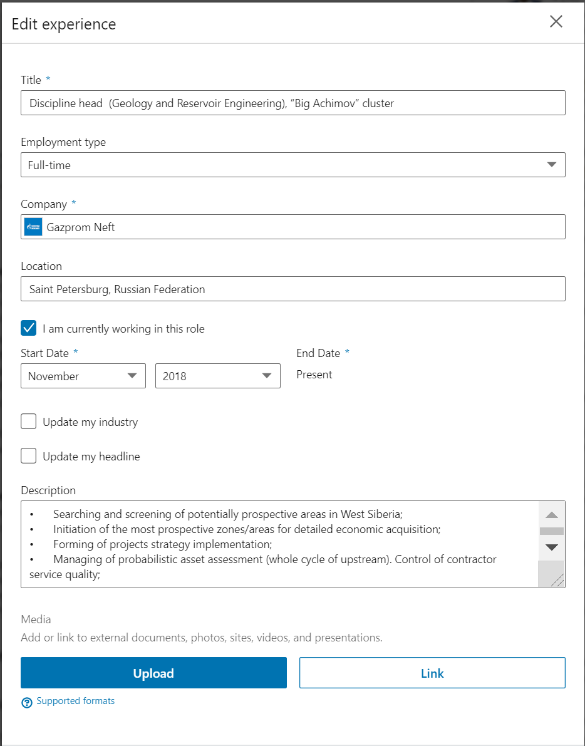


In edit contact info page pay attention on selection fields with fixed number of rows.

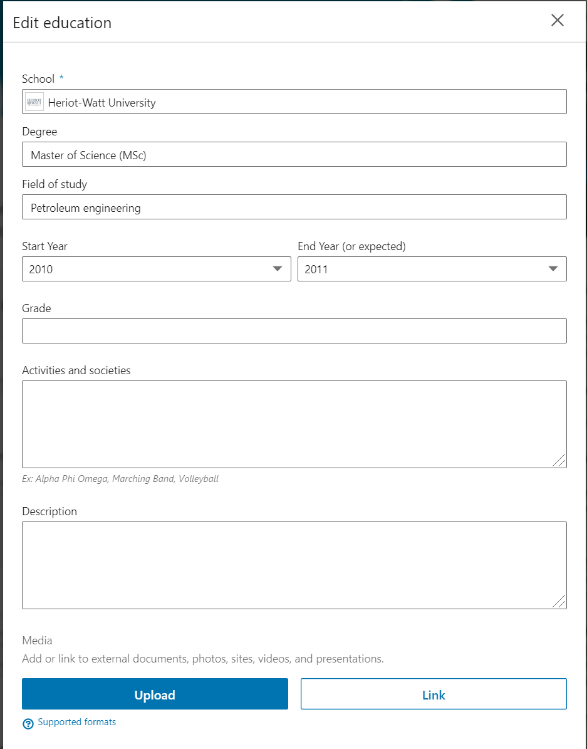
****

Profile edit button opens the menu with detailed profile structure. Thus, user can see required fields (marked by asterisks) and some fields with selection like “Current Position”.

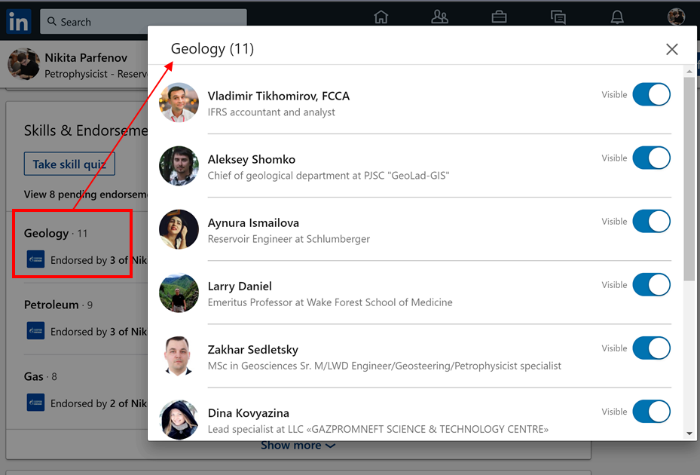
*\*To simplify the database structure location field and industry information were reflected in database as a TEXT fields (not selection fields like in the network).*

****

The work experience contains title, employment type (selection field), company name, location, period of work, description and allowance to add some media file or link. Supported media file types in this block are the next: .pdf, .ppt/.pptx, .doc/.docx, .jpg/.jpeg, .png, .gif

****

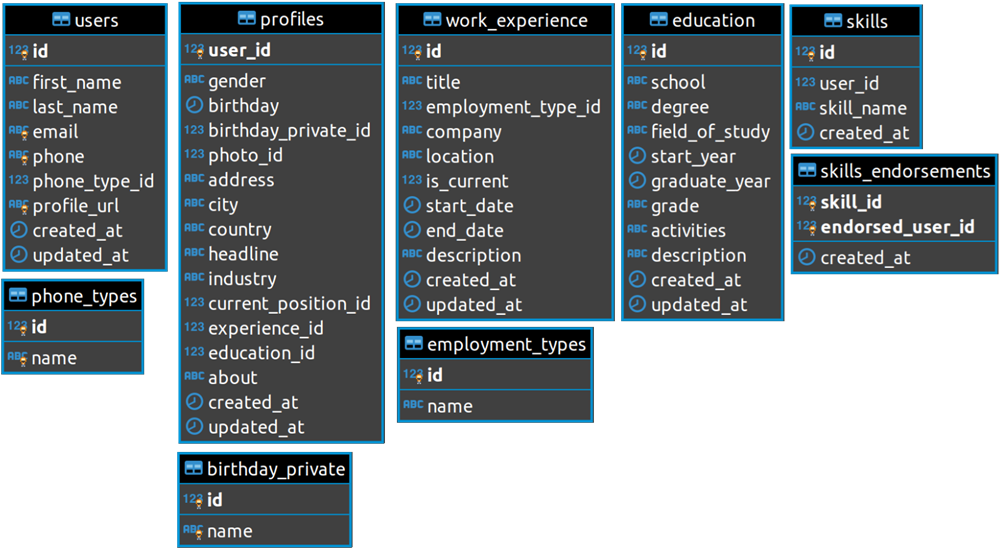
The education information has the similar structure as work experience. The difference in some text fields.



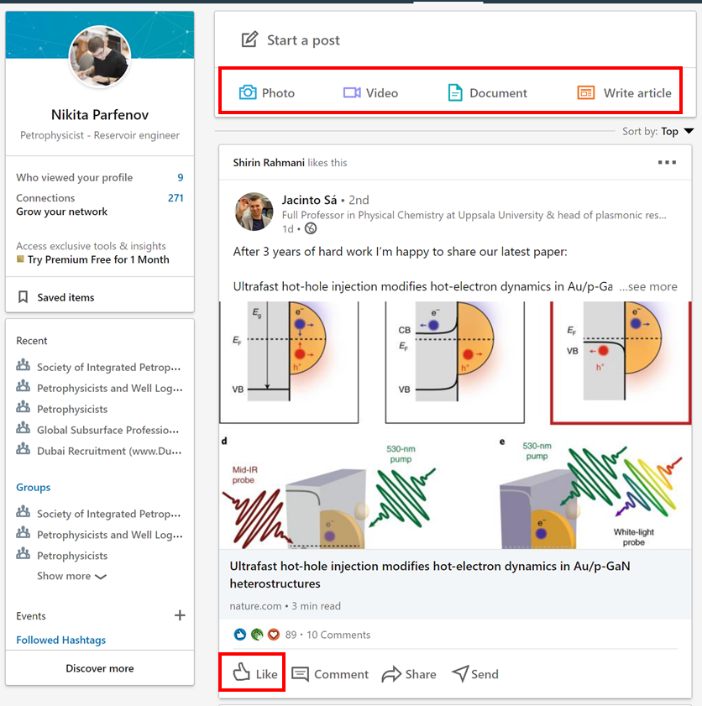
Skills & Endorsements block reflects “one to many” relationship where each skill chosen by user is endorsed by other users.

Note that the general users profile information has many fields which is weighted the table. In the same time users profile is the main object in social databases that is requested frequently. To optimize database loading the table of general user's information has been separated into users table (main user’s information) and profiles table (additional user’s information).

Users profile block tables are presented below.



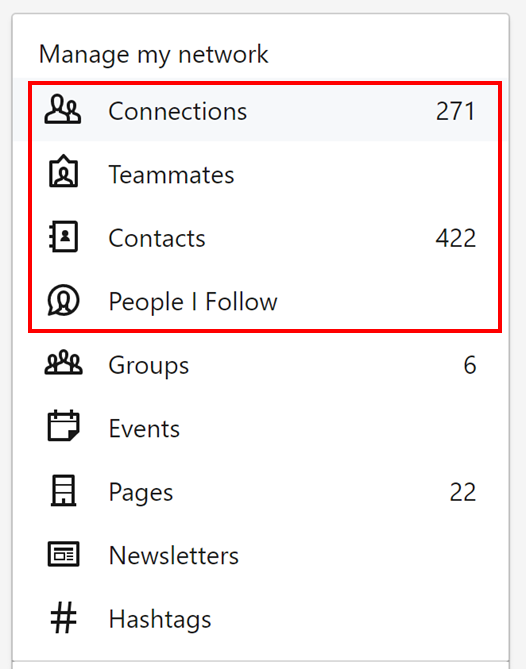
**2. Users social interaction**

****

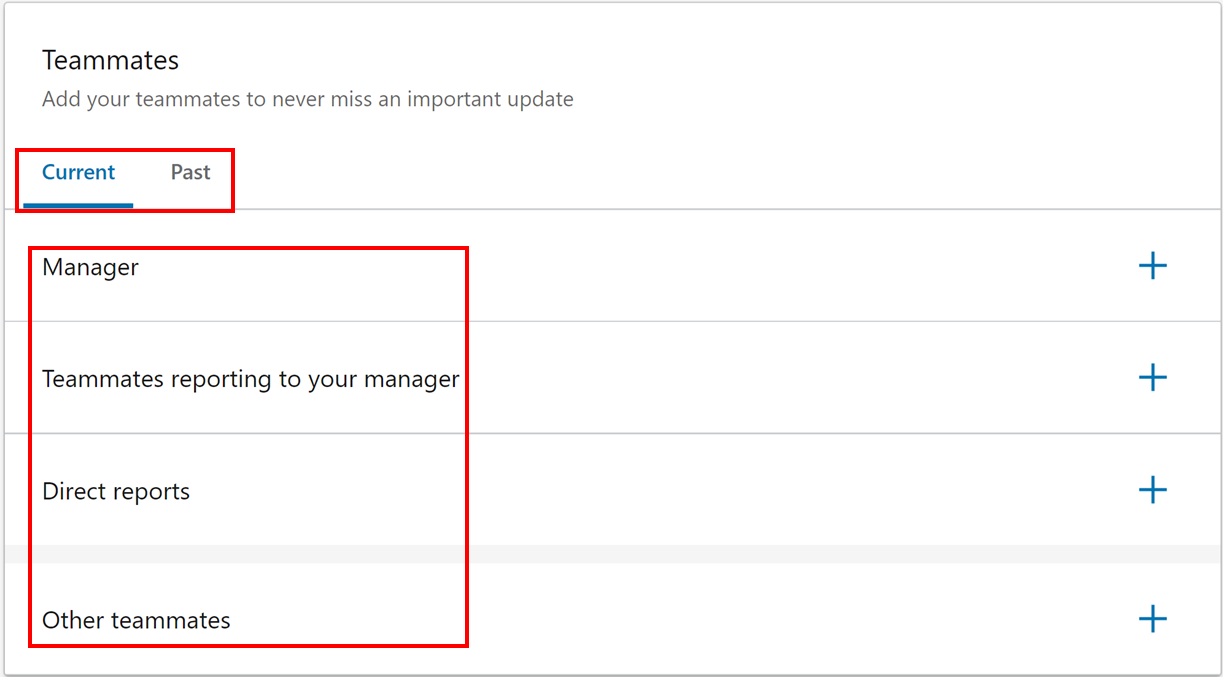
Each user can share posts with media content like photo, video, document or just article. Other users can like or comment posts.

Note that media content in LinkedIn is allowed only as a part of posts (exception – doc or image files in education and work experience pages) that is why it was related into social interaction block.

*\*To simplify the database comments to posts has been omitted.*

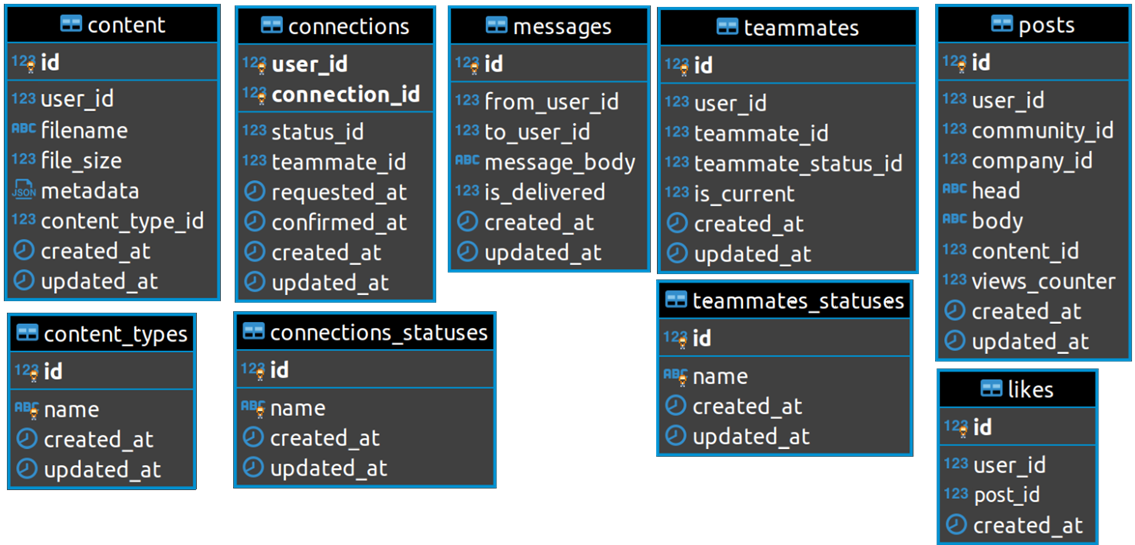
****

User’s network has 4 types of friendship: connections, teammates, contacts and people I follow. The 1st one if a general table of friendships, “contacts” and “people I follow” can be reflected as an additional flag – connection type.

****

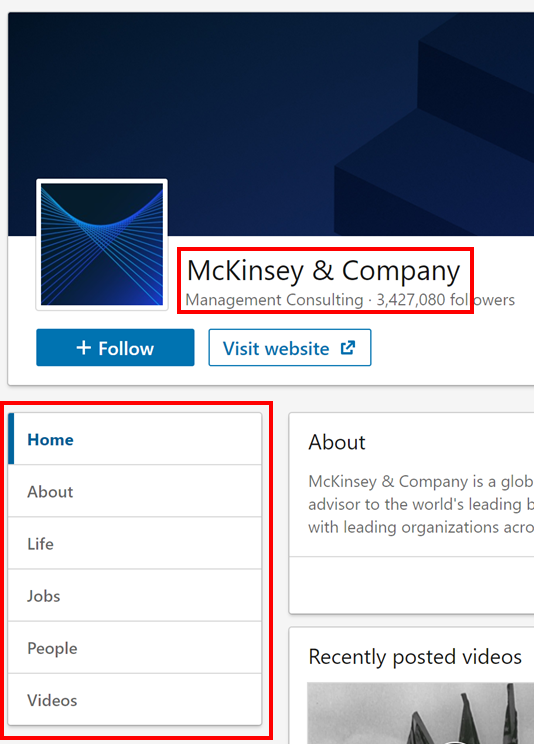
Teammates page has an interesting structure. It contains 4 types of relationships and 2 time flags.

Users social interaction block tables are presented below.

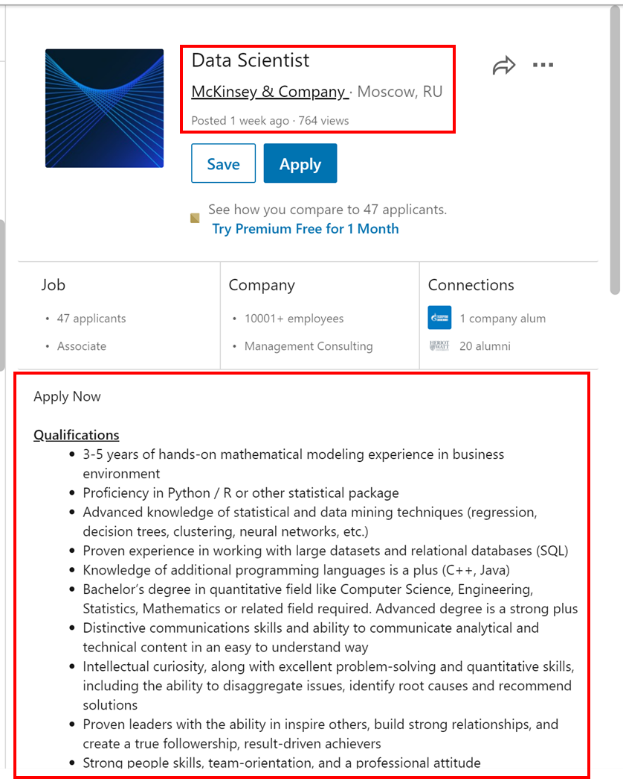
****

**3. Groups and Companies information**

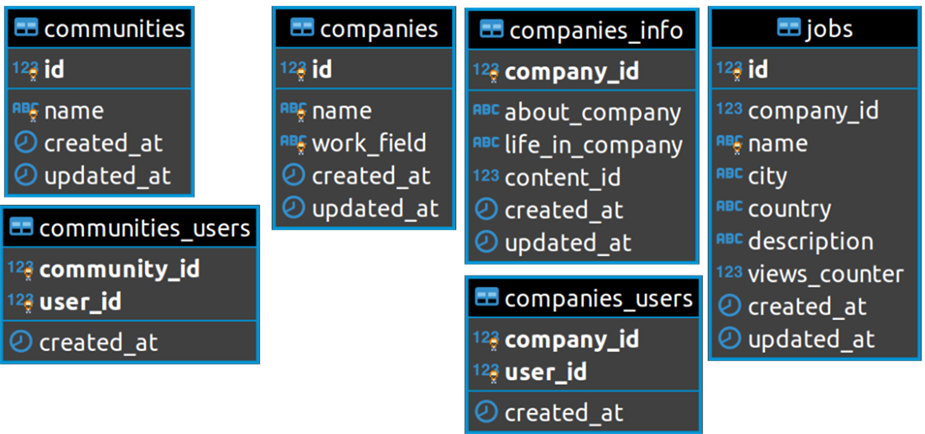
Groups in LinkedIn has a general “one to many” relationship type, so that was simplified up to groups name and connections between groups and users. Companies pages have more interesting structure, thereby it was considered in detail.



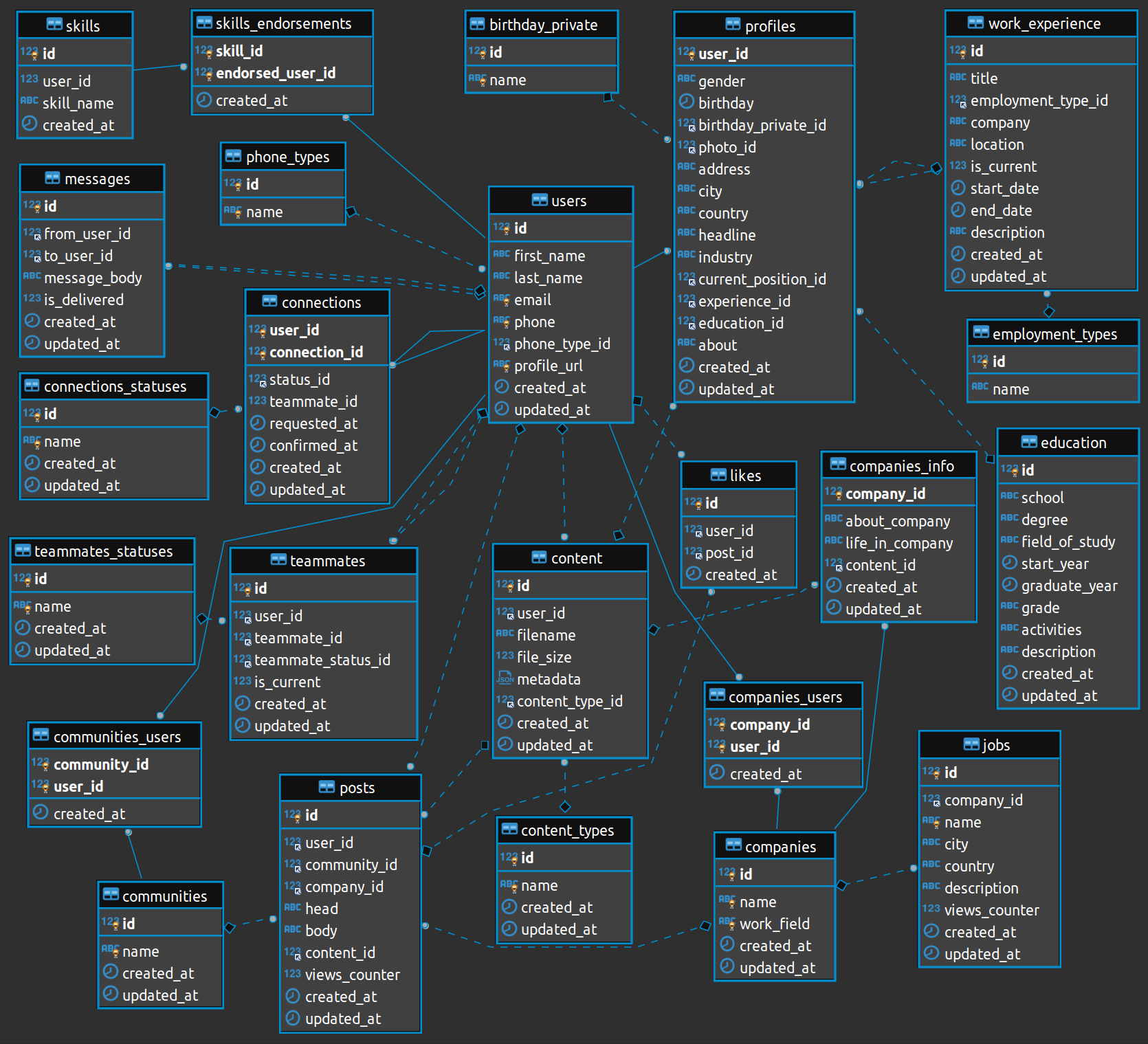
General company page has company name, type of business and pages with company information like “about”, “life” etc.



The main part of company page is a job pages which contains title, company name, location, posted time and description.

Groups and companies block tables are presented below.

**Entity-Relationship Diagram**

****

**Applications**

1. 1.1\_linkedin\_db-structure.sql
2. 1.2\_linkedin\_db-relations.sql
3. 1.3\_linkedin\_db-views\_and\_indexes.sql
4. 1.4\_linkedin\_db-log\_trigger.sql
5. 1.5\_linkedin\_db-fill\_by\_dummy\_data.sql
6. 2\_linkedin-queries.sql
7. Entity-Relation Diagram.png