

Database Systems: Final Project Part 2  
Kora S. Hughes & Michael Wang

## Homework Overflow

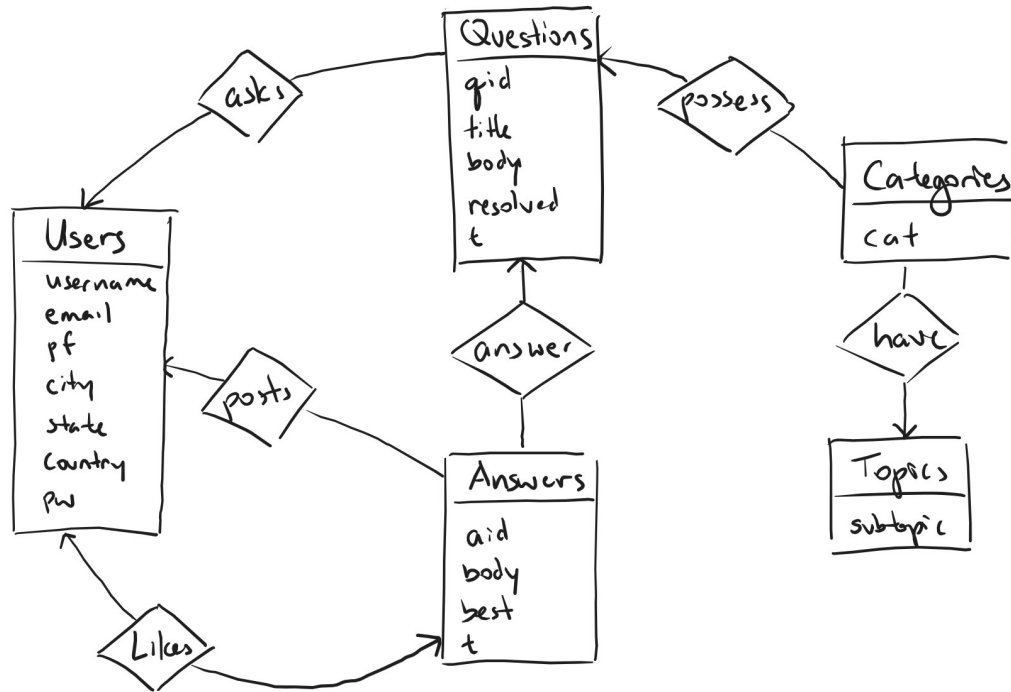
### System Design:

- System Overview:
  - Homework Overflow is a Q and A focused database system and website where students, professors, and any education enthusiasts can post questions and answers regarding their courses, homework, projects, and a plethora of subjects. Questions and answers on the site are visible by all visitors, but only users that sign up with an email, password, and username can post questions and answers.
- Code Overview:
  - The website of Homework Overflow is coded in HTML and PHP with a few UI components in CSS. Additionally, SQL is used to set up our backend data structures and communicate with the frontend interface via queries.
- How Users Interact with the System:
  - *Signup/Login/Logout Page:*
    - Users can generate a profile with a username, password, and email address, plus optional information like city, state, country, and a short profile description.
      - Once a user's profile is created, they can use their username to log in and out of their account.
  - *Post Page:*
    - Users can ask questions by providing a title and a detailed description of the inner workings of their question. Users are required to attach at least one (1) category to a question so other users have an easier time finding them. Once a user gets a response, they can like the answer and/or mark one as the best answer. If the user feels satisfied with the responses they get, they can then mark their question as resolved.
  - *Question/Answer Page:*
    - Additionally users can post answers to questions via the 'Post your answer' link at the bottom of each question page.
    - As users ask questions and post answers to other questions, their tier will increase so everybody can see how great of a scholar they are.
- Users can navigate through various questions and answers via 3 ways:
  - The first way is the *Browse* tab on the left of the website. Through this, users can click on various categories and subcategories to view a list of associated questions sorted by recency.

- The second way is to use the *Search* tab. This allows the user to input a series of keywords into the search bar as well as any number of categories clicked in the multi select search box. Once the submit button is pressed, the user will be shown a list of questions pertaining to each category the user chose (or any category if left blank) and the text in the search box. (specifics in the implementation section)
  - Note: each search box represents a component of a predefined 2 level hierarchy of categories, users can search an entire category or just the subcategory depending on which options are selected.
- The third and final way is to use the *Home* tab. This will simply show a list of the 10 most recent questions posted to Homework Overflow.
- *Profile Page*: Users can also view profiles. By clicking the username of any user, users can see the rank, profile information, and a comprehensive breakdown of all questions and answers posted in addition to all answers posted in response to each question.
  - If a user is logged in and is viewing his/her own profile, the email is also shown, along with the option to update the user's profile information.

**Backend Design:**

ER Diagram:

Schema: (**bold** = primary key, underlined = foreign key)

Users: contains user information	
<b>username</b>	varchar
Email	varchar
pf	text
City	varchar
State	varchar
Country	varchar
pw	varchar

Questions: contains question information	
<b>qid</b>	varchar
title	text
body	text
resolved	integer
t	datetime
username	varchar

username → references Users(username)

Answers: contains answers to questions	
<b>aid</b>	varchar
body	text
best	integer
t	datetime
<u>qid</u>	varchar
<u>username</u>	varchar

qid → references Questions(qid)

username → references Users(username)

Likes: maps who likes what answer	
<u><b>aid</b></u>	varchar
<u><b>username</b></u>	varchar

(aid + username) comprise the primary key for this table

aid → references Answers(aid)

username → references Users(username)

Categories: maps which questions are tagged under what categories/subcategories	
<b>cat</b>	predefined varchars

<b>qid</b>	varchar
------------	---------

(cat + qid) comprise the primary key for this table

qid → references Questions(qid)

<b>Topics</b> : maps which categories have subcategories	
<b>subcat</b>	predefined varchars
<u>cat</u>	predefined varchars

cat → references Categories(cat)

## Implementation:

### Database:

- As seen above, our database has tables for users, questions, and answers. All these take care of most of the heavy lifting in terms of data storage relating to how users interact with the system.
- Additionally, the likes table keeps track of which users like which answers, the category table keeps track of which question is attached to which categories (if any), and the Topics table keeps track of which subcategories (in this case 'subcategories') are attached to which supercategories.

### Source Code:

- Overall we use various SQL queries to select, update, insert, and delete elements of our database in a straightforward manner
- User logging instances are handled via sessions that are passed in between pages
- We connect our database to our website using a localhost instance (which can be found in `sql_connect()` in `functions.php`)

### SQL Integration: a general explanation of major queries used in our database

- Home page:
  - Query for all questions sorted by time descending

```
SELECT * FROM Questions JOIN Categories USING(qid) ORDER BY t
DESC
```

- Used PHP to display the first 10 questions and link to the question page, the browse page, and profile pages

- Browse page:
  - If no category is selected:
    - Display all top-level categories

```
SELECT cat FROM Topics WHERE cat=subcat
```

- If a top-level category is selected:
  - Query for all categories under the category \$cat

```
SELECT subcat FROM Topics WHERE cat='{ $cat }' AND
subcat!='{ $cat }'
```

- PHP displays these subcategories as links to them

- Query for questions under the top-level category \$cat ordered by time descending

```
SELECT *
FROM questions JOIN (
    SELECT qid
    FROM topics RIGHT JOIN categories ON
```

```

        topics.subcat = categories.cat
    WHERE topics.cat='{ $cat}'
    ORDER BY qid) AS c USING(qid)
ORDER BY t DESC

```

- PHP used to display the details of the question along with links to the appropriate pages

- If a sub-category is selected:

- Query for the top-level category

```
SELECT cat FROM Topics WHERE subcat='{ $cat}'
```

- Display the category for user convenience

- Query for questions under the subcategory \$cat ordered by time descending

```
SELECT * FROM Questions JOIN Categories USING(qid) WHERE
cat='{ $cat}' ORDER BY t DESC

```

- PHP used to display the details of the question along with links to the appropriate pages

- Search:

- Display multi-select box to choose categories to search within

- Query for top-level categories

```
SELECT cat FROM Topics WHERE cat=subcat
```

- Query for subcategories under each top-level category

```
SELECT subcat FROM Topics WHERE cat!=subcat AND
cat='{ $row["cat"]}'

```

- Create a temporary table storing keywords

- Create the table

```
CREATE TEMPORARY TABLE keywords(word varchar(24))
```

- Use a prepared statement to insert keywords from the front-end to the temporary table

```
INSERT INTO keywords (word) VALUES (?)
bind_param("s", $keyword)

```

- Search for keyword matches within the category chosen

```
SELECT qid, resolved, numA, title, t, categories.cat, username
FROM categories JOIN(
    SELECT resolved, c.username, c.t, c.qid, title,
        count(aid) AS numA
    FROM answers RIGHT JOIN (
        SELECT resolved, username, t, title, qid,

```

```

        count(qid) AS numQ
    FROM Questions JOIN Categories USING(qid),
        keywords
    WHERE LOCATE(keywords.word, questions.title) > 0
    GROUP BY qid)
    AS c ON Answers.qid = c.qid
GROUP BY c.qid
ORDER BY numQ desc, resolved desc, numA desc, t desc)
as d using(qid)
left join Topics on categories.cat = subcat
$categorySearch

//Where category search is
$categorySearch = "";
if (isset($_GET["category"])) {
    $categorySearch =
        "where subcat='" . implode("' or subcat='",
        $_GET["category"]) . "'";
}

```

- This search returns all questions with a matching keyword in the title and with a matching category
- The results are ordered on:
  1. The number of keyword matches
  2. If the question is resolved
  3. The number of questions asked
  4. Time posted
- User Profile Page:
  - we get the user's information to display via a given username
  - ```
"SELECT * FROM Users WHERE username='{$_u}'";
```
  - User Rank: Shows the user rank as text based on the number of posts a given user has made → 

```
"SELECT * FROM Posts";
```

    - Posts View: A common query for our system was figuring out the number of “posts” a user has made. We chose to define “posts” as the sum total of all questions and answers a user has created. So our SQL query involved



creating an itemized view of all users and their number of posts. This is primarily used in the calculation of a user's rank on the profile page.

```
$USER_TIERS = array("Beginner" => 4, "Intermediate" => 8, "Expert" => 10);

$sql = "SELECT * FROM Posts";
$result = $conn->query($sql);
if ($result->num_rows > 0) {
    while($row = $result->fetch_assoc()) {
        if ($u=$row['username']){
            if($row['num_posts'] < $USER_TIERS["Beginner"]){
                $rank = "Beginner";
            } else if($row['num_posts'] < $USER_TIERS["Intermediate"]){
                $rank = "Intermediate";
            } else {
                $rank = "Expert";
            }
        }
    }
}
```

```
create view Posts as
with all_posts as
((select username, count(*) as posts
from Questions
group by username)
union
(select username, count(*) as posts
from Answers
group by username))
select username, sum(posts) as num_posts
from all_posts
group by username;
```

- Then we get all the questions the user has posted:

```
"SELECT * FROM Questions join Categories using(qid) WHERE username='{ $u }'";
```

- From there, for each question we output its information and a list of the answers relating to said question with the like count under each answer

```
"SELECT title FROM Questions WHERE qid='{ $qid }'";
```

```
"SELECT * FROM Answers WHERE qid='{ $qid }' order by best desc, t desc";
```

```
"select count(aid) as num from likes join answers using(aid) where aid = '{ $row["aid"] }' group by aid";
```

- Then we show all the answers in response to any question that users have made and count the likes from that answer with the same query structure as above

```
"SELECT Answers.aid, Answers.body, Answers.t, Questions.title, Questions.qid FROM Answers join Questions using(qid) where Answers.username = '{ $u }'";
```

```
"select count(aid) as num from likes join answers using(aid) where aid = '{ $row["aid"] }' group by aid";
```

- Post Question Page:

- To post a question we generate a new question id by adding 1 to the current max

```
$sql = "SELECT MAX(qid) as qid FROM Questions";
$qid = strval( value: intval($conn->query($sql)->fetch_assoc()["qid"]) + 1);
```

- Then we add this question to our question table

```
"INSERT INTO Questions(qid, title, body, resolved, t, username) VALUES ( '{$qid}', '{$title}', '{$body}', 0, '{$t}', '{$_SESSION["user"]}' );"
```

- Then for each selected category we add its qid as an entry into the category table

```
"INSERT INTO Categories(cat, qid) VALUES (?, '{$qid}');"");
```

- Show Question/Answers Page:

- First we get the question and its categories

```
"SELECT * FROM Questions join Categories using(qid) WHERE qid='{$qid}'";
```

- Then we get all of its answers

```
"SELECT * FROM Answers WHERE qid='{$qid}' order by best desc, t desc";
```

- For each of those answers we also use the queries below to show/handle likes and selecting best answer

```
"SELECT Likes.username FROM Likes JOIN Answers USING(aid) WHERE aid='{$row["aid"]}";
```

```
"select count(aid) as num from likes join answers using(aid) where aid = '{$row["aid"]}' group by aid";
```

- Likes: helper to handle liking a question:

- We check if the user has liked this question

```
//if already liked, remove like, otherwise add like
$sql = "SELECT * FROM Likes WHERE aid='{$_GET["aid"]}' AND username='{$_SESSION["user"]}";
```

- If they have we like and if the haven't we remove the like (unlike)

```
if($result->num_rows > 0){
    //remove like from db
    $sql = "DELETE FROM Likes WHERE aid='{$_GET["aid"]}' AND username='{$_SESSION["user"]}";
} else {
    //add like to db
    $sql = "INSERT INTO Likes(aid, username) VALUES ( '{$_GET["aid"]}', '{$_SESSION["user"]}' );"
```

- Best: helper to mark an answer as the best (if the user owns this question)

- We get the question by qid

```
$sql = "SELECT * FROM Answers WHERE qid='{$_GET["qid"]}";
```

- Update the answer as best

```
$sql = "UPDATE Answers SET best='1' WHERE aid='{$_GET["aid"]}";
```

- Resolved: helper to handle marking questions as resolved

```

$sql = "SELECT resolved FROM Questions WHERE qid='{$_GET["qid"]}'";
$result = $conn->query($sql)->fetch_assoc();

//swap resolved boolean
if($result["resolved"] == 0){
    $sql = "UPDATE Questions SET resolved='1' WHERE qid='{$_GET["qid"]}'";
} else {
    $sql = "UPDATE Questions SET resolved='0' WHERE qid='{$_GET["qid"]}'";
}

```

- Answers Page:
  - Queries for the information of the question the user is responding to using \$qid as a previously passed question id variable (we use the username, body, t from this query to display and store information)

```
$sql = "SELECT * FROM Questions join Categories using(qid) WHERE qid='{$_GET["qid"]}'";
```

- We then add the answer to our table using this insertion query

```

//get new aid for this answer
$sql = "SELECT MAX(aid) as aid FROM Answers";
$aid = intval($conn->query($sql)->fetch_assoc()["aid"]) + 1;

//get time
$t = date("Y-m-d H:i:s");

//create new answer entry
$sql = "INSERT INTO Answers(aid, qid, body, best, t, username) VALUES ('{$aid}', '{$qid}', '{$answer}', 0, '{$t}', '{$_SESSION["user"]}');
$conn->query($sql);

```

- Login Page:
  - We check the user table for a matching username and password combination and this allows us to log the user in by storing their information in a session

```
"SELECT * FROM Users WHERE username='{$$_POST["username"]}' AND pw='{$$_POST["password"]}';
```

- Create Profile (register) Page:
  - First we query for the username to make sure its not taken

```
"SELECT * FROM Users WHERE username='{$$_POST["username"]}'";
```

- Then we insert the record and corresponding information into the table and use similar queries as that above to check if the user was inserted correctly

```

$sql = "INSERT INTO Users(username, pw, email, city, state, country, pf) VALUES
('{$username}', '{$pw}', '{$email}', '{$city}', '{$state}', '{$country}', '{$profile}');

```

- Edit Profile Page:

- We query for the user's information as default

```
"SELECT * FROM Users WHERE username='{$_GET["u"]}';
```

- And then we call many update queries depending on which the user input for email, username, city, country, state, and profile description.

```
- "UPDATE Users SET email='{$email}' WHERE username='{$_GET["u"]}'";
```

## Test Data:

| Users       |               |                     |             |       |         |                                                  |
|-------------|---------------|---------------------|-------------|-------|---------|--------------------------------------------------|
| username    | pw            | email               | city        | state | country | pf                                               |
| elmo        | cookiemonster | seseamest@gmail.edu | Brooklyn    | NY    | USA     | who even a                                       |
| john        | password      | johnsmith@nyu.edu   | Bronx       | NY    | USA     | just john smithin it up                          |
| kora        | 123           | khughes@nyu.edu     | Brooklyn    | NY    | USA     | help with hw plz i am ignorant<br>colleg student |
| michael     | 12321         | michaelwang@nyu.edu | New York    | NY    | USA     | how do you code?                                 |
| prof. lando | 345           | profussy@nyu.edu    | Springfield | CL    | USA     | professor tryna catch students<br>plagerizing    |

| Categories       |     |
|------------------|-----|
| cat              | qid |
| Math             | 1   |
| Science          | 2   |
| Math             | 3   |
| Math             | 4   |
| Foreign Language | 5   |
| History          | 6   |
| Spanish          | 6   |
| Economics        | 7   |
| Chemistry        | 8   |
| Computer Science | 9   |
| Chemistry        | 9   |
| Literature       | 10  |
| Oral History     | 11  |
| Math             | 13  |
| Math             | 14  |

| Questions |
|-----------|
|-----------|

| qid | title                                                | body                                                                                                             | resolved | t               | username    |
|-----|------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|----------|-----------------|-------------|
| 1   | Quic Mafs                                            | What is 1+1?                                                                                                     | 1        | 4/12/2022 18:59 | kora        |
| 2   | Human Disection                                      | how do you dissect a human?                                                                                      | 0        | 4/12/2022 18:11 | kora        |
| 3   | Hard Mafs                                            | What is the integral of $\cotan^4(x)$ ?                                                                          | 0        | 5/12/2022 12:41 | kora        |
| 4   | Middleschool curriculum: integrals???                | A middleschooler of mine turned in a complex integral dissertaion instead of his algebra homework, what do I do? | 0        | 3/12/2022 8:59  | prof. lando |
| 5   | How to be a polyglot?                                | I want to know all the languages but I only know english tehe - advice?                                          | 0        | 4/12/2021 10:23 | prof. lando |
| 6   | What is the the lost history of Spanish colonization | Mas sorces en espanol por favor?                                                                                 | 0        | 4/12/2019 15:48 | michael     |
| 7   | How to maximize capital gains tax for the rich?      | What even is capital gains?                                                                                      | 0        | 4/12/2020 12:15 | michael     |
| 8   | How to dissolve a body?                              | What is the optimal strong acid/base and how do I get rid of the waste wihtout anyone knowing?                   | 0        | 4/12/2021 17:36 | michael     |
| 9   | Quantum Superposition                                | Why does quantum computing use superposition when subatomic read/writes are costly?                              | 0        | 4/12/2021 20:51 | john        |
| 10  | This book changed my life                            | Haha gottem, you read? NEEEEERD                                                                                  | 0        | 4/12/2019 20:01 | john        |
| 11  | My grandma is a secret spy?                          | she told me that her ancestors were passed down from royalty to exterminate the Queen? What do I do?             | 0        | 4/12/2022 14:13 | elmo        |
| 13  | Math is hard                                         | what is math??? why do we need to take maf                                                                       | 0        | 4/12/2022 14:13 | kora        |
| 14  | Math                                                 | test                                                                                                             | 0        | 4/12/2022 14:13 | kora        |

| Topics           |                  |
|------------------|------------------|
| cat              | subcat           |
| English          | English          |
| English          | Literature       |
| Foreign Language | Afrikans         |
| Foreign Language | Arabic           |
| Foreign Language | Chinese          |
| Foreign Language | Foreign Language |
| Foreign Language | French           |
| Foreign Language | Russian          |
| Foreign Language | Spanish          |
| History          | History          |
| History          | Oral History     |
| History          | US History       |
| Math             | Computer Science |
| Math             | Economics        |
| Math             | Math             |
| Science          | Biology          |
| Science          | Chemistry        |
| Science          | Physics          |
| Science          | Science          |

| Answers |     |             |      |                 |          |
|---------|-----|-------------|------|-----------------|----------|
| aid     | qid | body        | best | t               | username |
| 1       | 1   | 1+1=2 doy   | 1    | 4/13/2022 14:11 | kora     |
| 2       | 2   | just dont   | 1    | 4/16/2022 19:12 | elmo     |
| 3       | 3   | 1+1=2 doy   | 1    | 4/19/2022 17:13 | elmo     |
| 4       | 1   | 1+1=21 easy | 0    | 4/18/2022 13:01 | elmo     |

|    |    |                                        |   |                 |             |
|----|----|----------------------------------------|---|-----------------|-------------|
| 5  | 4  | give them an A+++                      | 1 | 4/20/2022 11:05 | michael     |
| 6  | 7  | invest in crypto                       | 1 | 4/21/2022 12:32 | michael     |
| 7  | 10 | rude                                   | 1 | 4/22/2022 8:53  | michael     |
| 8  | 11 | tap into the super secret spy network  | 1 | 4/23/2022 10:12 | kora        |
| 9  | 1  | 1+1=10 binary right?                   | 0 | 4/21/2022 2:11  | michael     |
| 10 | 1  | Are you actually asking this right now | 0 | 4/14/2022 14:11 | prof. lando |
| 11 | 13 | Math is important                      | 0 | 4/14/2022 14:11 | prof. lando |

| Likes |             |
|-------|-------------|
| aid   | username    |
| 2     | elmo        |
| 3     | elmo        |
| 1     | john        |
| 4     | john        |
| 1     | kora        |
| 2     | michael     |
| 3     | michael     |
| 5     | prof. lando |

- Logged Sessions:
  - We use PHP sessions to track if the user was logged in. It is empty if the user is not logged in and is set to the username of the user if logged in.
    - This variable is used to apply restrictions to likes, marking questions as resolved, marking best answers, and editing profile information.
      - The user must be logged in to the account that posted a question to mark the best answer to a question or mark it as resolved
      - If the user is viewing his/her own account information, the email is displayed and so is an option to edit profile information



### Highlights For Extra Credit:

- Source Code Highlights:
  - Depending on where the user searches we optimize our website's ability to sort and rank questions and answers situationally based on factors such as time, user interactions with a question, whether or not a question is resolved, whether an answer is marked as the best, number of likes, and rank of a user.
  - Cross-Site Scripting is prevented by using PHP's built-in function 'htmlspecialchars()' which replaces any characters significant to HTML with non-significant characters
    - This function is used whenever we pass data in, whether using PHP's GET or POST
  - In order to accommodate users, we build our system such that information for searches and category browsing are passed through the url so users can freely save and bookmark a question they like if they wish to save it later for studying.
  - We created a function to log to the console to assist in debugging
- Front End of Note:
  - Easy access side bar and comfortable UI using CSS
  - Users can edit their profile information
  - Users can mark a best answer on questions they've created
  - Our search page utilizes a special task menu of categories and subcategories users can use to filter their searches alongside keywords, emphasizing functionality in keyword-category query combinations and shows the user how hierarchies of categories are dynamically outlined.