# Database Design

* An experiment is an ordered set of questions.
* When user finishes going through the experiment's questions, the responses are associated to a trial
* I have attempted to [normalize](https://en.wikipedia.org/wiki/Database_normalization) database design, which theoretically makes it easier to ensure data integrity and reduces data redundancy.

### Legend:

* (K): Primary key, are unique and non-empty values that identify a row across tables
* (U): Column will have unique non-empty values
* (NE): Non-empty values, can be repeating
* None: Can have repeating or empty values

## Tables linking volunteers to their questions

**users**

* New users cannot be added until they have a unique IP Address

|  |  |
| --- | --- |
| worker\_id | TEXT (U) (K) |
| last\_name | TEXT |
| first\_name | TEXT |
| password\_hash | LONG |
| ip\_address | TEXT (U) |

**users\_experiments**

* Trial ID has -1 if the user has not yet finished submitting the responses

|  |  |
| --- | --- |
| worker\_id | TEXT (K) |
| experiment\_id | INTEGER (K) |
| trial\_id | INTEGER |

**experiments\_questions**

* Experiment ID has one-to-many relation with Question ID.
* Questions will be shown in the increasing values of Question order.
* Questions with the same order will be displayed randomly.

|  |  |
| --- | --- |
| experiment\_id | INTEGER (K) |
| question\_order | INTEGER |
| question\_id | INTEGER (K) |

## Experiment data collected

**responses**

|  |  |
| --- | --- |
| trial\_id | INTEGER (K) |
| question\_id | INTEGER (K) |
| response\_time | TIME |
| response | TEXT |

## Questions provided by experimenters

**questions**

* Question type tells us which table we should look in to find more about that question ID.
* Question type is going to be MCQ, Text Response, Image Reaction…

|  |  |
| --- | --- |
| question\_id | INTEGER (K) |
| question\_type | CHAR |

**mcq**

* Assuming all multiple-choice questions will have a maximum of 4 choice.
* Having NULL as a choice means there are no more choices, suitable for framing true/false questions out of this model

|  |  |
| --- | --- |
| question\_id | INTEGER (K) |
| question | TEXT |
| choice\_1 | TEXT |
| choice\_2 | TEXT |
| choice\_3 | TEXT |
| choice\_4 | TEXT |

**image**

* Regarding questions where the user has to type after seeing an image
* Image will be displayed for the set\_for time and we would wait for max\_time for the user to response

|  |  |
| --- | --- |
| question\_id | INTEGER (K) |
| question | TEXT |
| image\_type | TEXT |
| show\_for | TIME |
| max\_time | TIME |

**text**

* Having NULL as hint means there no hint needs to be displayed where the user has to type the answer

|  |  |
| --- | --- |
| question\_id | INTEGER (K) |
| question | TEXT |
| hint | TEXT |