

Prof R. Madana Mohana



BIG DATA ANALYTICS

MongoDB

Data Modelling

<https://www.youtube.com/c/RASINENIMADANAMOHANA>

MongoDB: Data Modelling

MongoDB: *Data Modelling*

- **Data** in **MongoDB** has a flexible **schema.documents** in the same **collection** (Table).
- They do not need to have the same set of **fields** or **structure**, and **common fields** in a **collection's documents** may hold **different types** of data.

MongoDB: *Data Modelling*

Some considerations while designing Schema in MongoDB:

- *Design* your **schema** according to **user requirements**.
- *Combine* objects into one document if you will use them together. Otherwise separate them (but make sure there should not be need of **joins**).
- *Duplicate* the data (but limited) because **disk space** is cheap as compare to compute time.
- *Do* joins while **write**, not on **read**.
- *Optimize* your schema for most frequent **use cases**.
- *Do* complex **aggregation** in the schema.

MongoDB: *Data Modelling*

Example:

Suppose a client needs a **database design** for his blog/website and see the differences between **RDBMS** and **MongoDB** schema design.

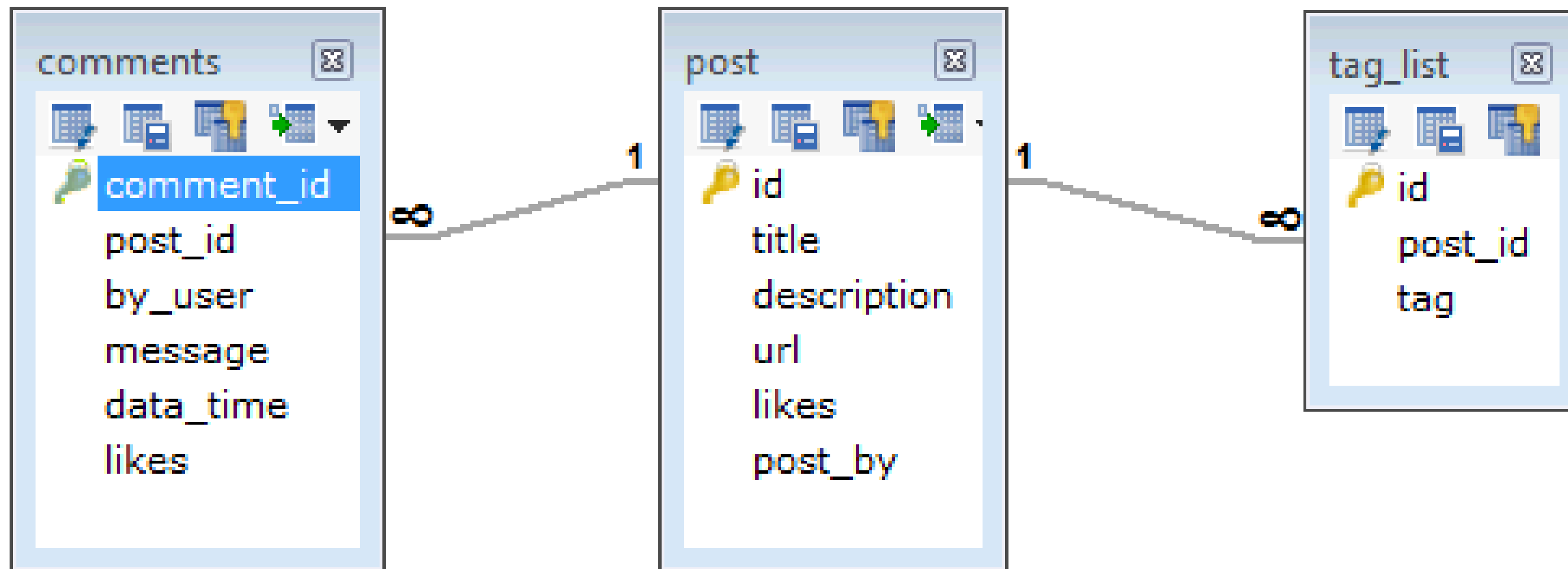
Website has the following requirements.

- Every post has the **unique title, description** and **URL**.
- Every post can have one or more **tags**.
- Every post has the name of its **publisher** and total number of **likes**.
- Every post has **comments** given by **users** along with their **name, message, data-time** and **likes**.
- On each post, there can be **zero** or **more comments**.

MongoDB: *Data Modelling*

Example:

In **RDBMS** schema, design for above requirements will have *minimum three tables*.



MongoDB: *Data Modelling*

While in MongoDB schema, design will have one collection post and the following structure:

```
{  
  _id: POST_ID  
  title: TITLE_OF_POST,  
  description: POST_DESCRIPTION,  
  by: POST_BY,  
  url: URL_OF_POST,  
  tags: [TAG1, TAG2, TAG3],  
  likes: TOTAL_LIKES,  
}
```

MongoDB: *Data Modelling*

While in MongoDB schema, design will have one collection post and the following structure:

```
comments: [  
{  
  user: 'COMMENT_BY',  
  message: TEXT,  
  dateCreated: DATE_TIME,  
  like: LIKES  
},
```


MongoDB: *Data Modelling*

While in MongoDB schema, design will have one collection post and the following structure:

```
{  
user: 'COMMENT_BY',  
message: TEXT,  
dateCreated: DATE_TIME,  
like: LIKES  
}  
]  
}
```

MongoDB: *Data Modelling*

So while showing the data, in RDBMS you need to join three tables and in MongoDB, data will be shown from one collection only.