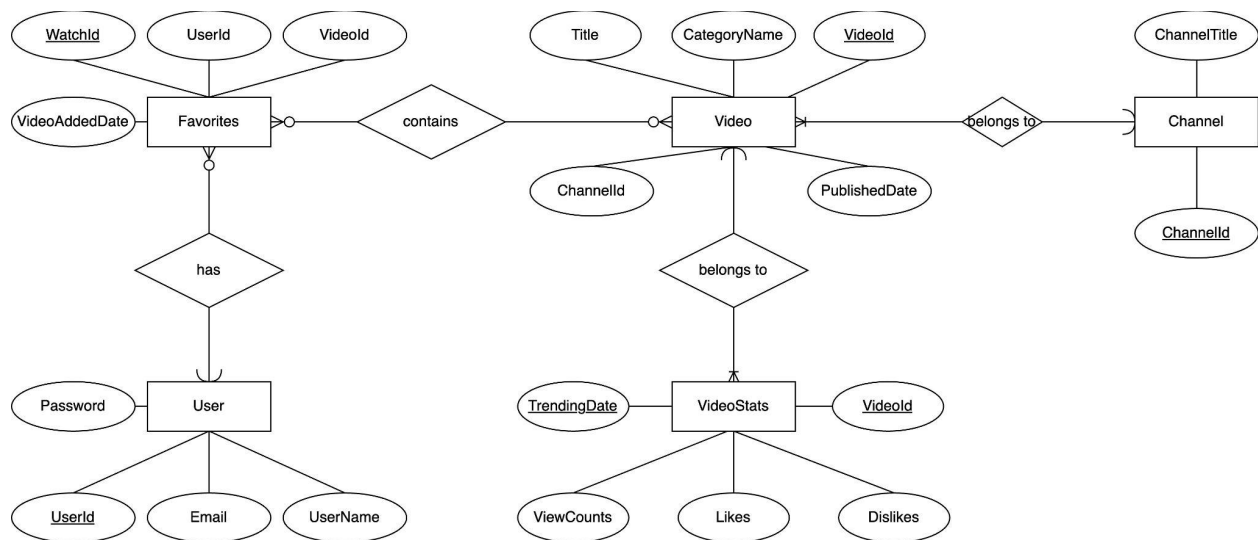


## ER Diagram



## Relational Schema Definition

```

User(
    UserId:CHAR(6) [PK],
    UserName:VARCHAR(100),
    Email:VARCHAR(255),
    Password:VARCHAR(255)
)

Favorites(
    WatchId:CHAR(8)(PK),
    UserId:CHAR(6),
    VideoId:CHAR(11),
    VideoAddedDate: VARCHAR(50)
)

Video(
    VideoId:CHAR(11)[PK],
    Title:VARCHAR(100),
    PublishedDate: VARCHAR(50),
    CategoryName:VARCHAR(100),
    ChannelId:CHAR(24) [FK to Channel.ChannelId]
)
    
```

```
VideoStats(
    VideoId:CHAR(11) [PK, FK References to Video.VideoId],
    TrendingDate: VARCHAR(50)[PK],
    ViewCount: INT,
    Likes: INT,
    Dislikes: INT
)
```

```
Channel(
    ChannelId:CHAR(24) [PK],
    ChannelTitle:VARCHAR(255)
)
```

## Description and Assumptions of Relations:

In general, we aim to develop a hot video recommendation application based on the dataset of daily trending YouTube Videos. The above relational scheme consists of five tables: User, Favorites, Video, and Category, as well as Channel. We aligned our schema data types using the dataset listed types and the standard Youtube character regulation limitations.

- **User**
  - **Description:** The User table stores user information, including their unique UserId, UserName, Email, and Password. The UserId serves as the primary key of the User table.
  - **Assumption:** We proposed that every user could have 0 to many favorite lists and one favorite list can only be associated with one user.
- **Favorite**
  - **Description:** The Favorites table stores user favorites, with each record consisting of a unique WatchId, the corresponding UserId from the User table (via a foreign key), the VideoId of the favorite video, and the date the video was added. The WatchId serves as the primary key of the Favorites table.
  - **Assumption:** Each favorite list may contain 0 or many videos, and a single video can be liked by many users, thus one video can belong to 0 or many favorite lists.
- **Video**
  - **Description:** The Video table stores information about videos, including their unique identifier "VideoId", the video's title "Title", the date the video was published "PublishedDate", and the name of the category "CategoryName". The "VideoId" is the primary key of the table, which means it uniquely identifies each record in the table. The "CategoryName" column represents the category that the video belongs to.
  - **Assumption:** Further analysis of the dataset revealed that each video has one or multiple records of video stats, but one record can only identify one movie.
- **Channel**
  - **Description:** The Channel table stores information about each video channel, including its unique ChannelId, the ChannelTitle. The ChannelId serves as the primary key of the Channel table.

- **Assumption:** Each video belongs to only one channel, but a channel can contain 1 to many videos.
- **VideoStats**
  - **Description:** The "VideoStats" table stores statistical information for videos, including the video ID, trending date, video view counts, likes, and dislikes. The VideoId and TrendingDate are the primary key for the table, and the video ID is a foreign key referencing the "Video" table. The table is designed to track the number of views, likes, and dislikes for each video on specific trending dates.
  - **Assumption:** As previously mentioned, each record in the "VideoStats" table represents the statistics for one video on a specific trending date, and a single video can have multiple records in the table.