

HW1: (E)ER diagram construction

In this assignment, you need to create a **conceptual (Extended) Entity-Relational [ie. (E)ER] diagram**, to model the scenario described below; note that your design is not going to be at an implementation level, ie. you don't have to worry about representing your design using relational tables (including bridges) - you will only need to come up with **entities** and **relationships**, and possibly also use the (E)ER notation for supertypes/subtypes.

Please submit your work as a single image file in .jpg or .png form that shows the entire diagram, via D2L Dropbox (not via Blackboard). ALSO include a README.txt that contains any design choices you want to highlight, and/or assumptions you made [if the (E)ER diagram were 'code', this would be 'comments'].

You can create the (E)ER diagram using any software of your choice, including:

- yEd (<http://www.yworks.com/products/yed>)
- draw.io (browser-based)
- Lucidchart (browser-based)
- Vertabelo (also entirely online)
- DIA
- Project Mogwai (<https://github.com/mirkosertic/MogwaiERDesignerNG>)
- E/R Assistant (Windows only:
https://highered.mheducation.com/sites/0072942207/student_view0/e_r_assistant.html)
- Visio
- <https://erdplus.com/standalone> - browser-based :)
- ...

After constructing the (E)ER diagram, save, or take a screengrab snapshot, submit it [as a .jpg or .png image file].

Note that you can even draw your diagram **legibly** on paper and take a photo of it and submit that - but having said that, I'd encourage you to use a diagramming tool, that will make your result look professional, and have you follow industry practice.

You need to use Crow's Foot Notation for the (E)ER diagram. For each relationship, indicate the cardinality (minimum and maximum participation), also via Crow's Foot symbols - use [this infographic](http://www.vivekmchawla.com/) (from www.vivekmchawla.com/) as a guide [you don't need to denote cardinality as (1,1) etc., instead, you would use the notation shown in the infographic, ie. use symbols such as |O and ||].

How much detail should your diagram contain? Use [this sample](#) as a guide (eg. you do not need to indicate data types for attributes).

Description [created by Vinutha Ananthachandran, one of our CPs :)]

YouTube is one of the most popular video-sharing platforms that allows its users to upload, watch, search and share video-based content. It is a complex and highly scalable system that mainly depends on its users to post videos to grow and earn profits. A small portion of this system can be described using the following components.

A user is identified with a unique id, name, email, age, address, etc. YouTube users can be video creators or video consumers or both. YouTube channels are like dashboards for users to upload and distribute their unique videos. Channels have a name, owner, subscription count and the details of when it was created. For simplicity, it is safe to assume that a channel is created by only one video creator but a video creator can have multiple channels. Video consumers can subscribe to 'N' number of channels; the details of which is available on both ends (subscribers, creators). Subscription can either be paid or free, and is identified via a subscription type.

YouTube videos can be categorized into informational videos, entertainment videos, etc. Informational videos are identified based on keywords and entertainment videos are identified based on tags. When a user uploads a video, YouTube's transcoding server starts the video encoding process before streaming it on to its platform. After

the completion of the encoding process, the video is stored, and the video metadata database is updated with the video details. This data includes a URL, title, thumbnails, category, duration, description, uploader id, upload date, and upload time.

YouTube content creators earn their revenues from the platform based on the popularity of their channels and videos. Each video has statistics to gauge popularity. YouTube statistics data includes likes, dislikes, view count, number of shares and number of comments. Each video can have 'N' number of comments in the comments section where users post their thoughts. The comments data includes video id, user id, comment text, likes, sentiment, and the details of when it was commented. As an optional source of income, YouTubers can use a third-party sponsor for their videos. For each video the sponsor details include the sponsor id, name, phone, address and the amount sponsored.

Your task is to create an (E)ER diagram that accurately represents the various entities and their relationships described in the above specification. You can also make some reasonable assumptions (which you'd document in your README) while building your (E)ERD.

Note that there isn't a single correct/perfect design, or even a single 'good' design - any design involves tradeoffs which determine their efficacy/viability. You are free to make intelligent choices about what data (attributes) to store where (entities), and how to connect all the pieces (relations). **Note too that some requirements stated above, might not be able to be captured in an (E)ER diagram - this is fine.** Be sure to document your design decisions (they would serve to provide rationale for "why you created your design the way you did").

Please do not plagiarize, 'work together', etc. If we see that your solution resembles anything else, we will need to report you to SJACS, etc. **How to avoid this? Simply do your own work!** This is a design problem, so you have much latitude in coming up with one that works - enjoy the process!

Submission checklist:

- .jpg or .png pic of your (E)ER diagram
- README.txt description file

You can post questions (and answer others' :)) on Piazza, under 'hw1'. **HAVE FUN!**
