

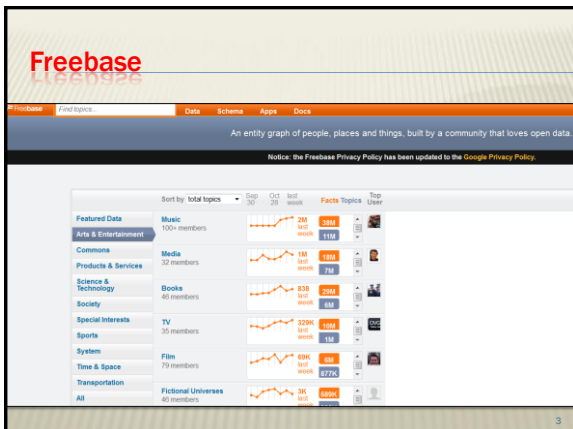
DB Project

Database Systems
Fall 2012-2013

Database project – TV



✕ <http://www.freebase.com/view/tv>



Freebase

Freebase | [Join topics](#) | [Data](#) | [Schema](#) | [Apps](#) | [Docs](#)


An entity graph of people, places and things, built by a community that loves open data.

Notice: the Freebase Privacy Policy has been updated to the Google Privacy Policy.

Featured Data	Sort by: Total topics	Step	Out	In	Facts	Topics	Top
Music	100+ members	2M	100K	100K	100K	100K	
Arts & Entertainment	32 members	1M	100K	100K	100K	100K	
Commons	Media	32 members	1M	100K	100K	100K	
Products & Services	Books	40 members	818	100K	100K	100K	
Science & Technology	Society	40 members	818	100K	100K	100K	
Special Interests	TV	30 members	329K	100K	100K	100K	
Sports	Film	79 members	89K	100K	100K	100K	
System	Time & Space	79 members	89K	100K	100K	100K	
Transportation	Fictional Universes	40 members	818	100K	100K	100K	
All							

What is Freebase

- ✕ Freebase is an open repository of structured data of almost 23 million entities.
- ✕ An **entity** is a single person, place, or thing. Freebase connects entities together as a graph.
- ✕ <http://www.freebase.com/>



What is Freebase (2)

- ✕ Metaweb
http://www.youtube.com/watch?v=TjfrNo3Z-DU&feature=player_embedded
- ✕ Freebase
http://wiki.freebase.com/wiki/What_is_Freebase%3F



Into Freebase (1)

Michael Richards	Michael Anthony Richards (born July 24, 1948) is an American actor, comedian, writer and television producer, widely known for his portrayal of Cosmo Kramer on the television sitcom Seinfeld. During the show's run, he received the Primetime Emmy.	Seinfeld Fridays Madhead Manor The Michael Richards Show	Cosmo Kramer Rick Vic Nardella Peterson
Julia Louis-Dreyfus	Julia Scarlett Elizabeth Louis-Dreyfus (née "Seidman"; born January 13, 1961) is an American actress, comedienne and producer, well known for her roles in Seinfeld, The New Adventures of Old Christine, and her current series Veep. She is also...	Seinfeld Madhead Manor The New Adventures of Old Christine Days of Our Lives The Art of Being Nick	Elaine Benes Eleanor Elser Riggs Christine Campbell Elaine Seltz Rachel
Jerry Seinfeld	Jimmie Allen "Jerry" Seinfeld (born April 29, 1954) is an American stand-up comedian, actor, writer, and television and film producer, best known for playing a semi-fictional version of himself in the sitcom Seinfeld (1989–1998), which he co-created.	Seinfeld One Night Stand The Marriage Ref	Jerry Seinfeld Frances Cosmo-Kramer Charles Pincus
Lauren Graham	Lauren Helen Graham (born March 16, 1967) is an American actress and producer. She is best known for playing Louise Platter on the 1993 Fox sitcom Aladdin. Other notable roles include Sarah Brannaman on Parenthood, Graham was born in Honolulu, Hawaii. Her...	Seinfeld Glee Lust for Life	Louise Platter Sarah "Sue" Seinfeld Andrea

Into Freebase (2)

Concord

Concord is a town in Middlesex County, Massachusetts, in the United States. As of the 2010 census, the town population was 17,668. The United States Census Bureau considers Concord part of Greater Boston. Concord has a large presence in American history and literature. The area which became the town of Concord was originally known as "Masketquad", situated at the confluence of the Sudbury and Assabet rivers. Native Americans had cultivated...

W Read article at Wikipedia

Date founded: 1635
Time zone(s): North American Eastern Time Zone
Area: 67,08909 km² (25.9 mi²)
Population: 16,993 (2000)
Also known as: Concord, Massachusetts

Place of birth
Country of birth
Gender: 16,993,0100000000002
Profession: 16,993,005

Into Freebase(3)

Freebase **boston** **Data** **Schema** **Apps** **Docs**

Select an item from the list:

- Boston** City/Town/Village
- Atlanta Braves** Baseball Team
- Boston** City/Town/Village
- The Boston Globe** Newspaper
- Boston** Band
- Boston Red Sox** Baseball Team
- Boston Symphony Orchestra** Conducted Ensemble
- Boston Herald** Newspaper
- South Boston** Neighborhood
- Boston University** College/University

view more

Science & Technology Film 79 items

Society People

Special Interests People

Boston

Musical Genres: Classic rock, Hard rock, Rock music

Place: Musical Career

Began: Boston

Active as Musical Artist (start): 1975

Boston is an American rock band from Boston, Massachusetts that achieved its most notable successes during the 1970s and 1980s. Centered on guitarist, keyboardist, songwriter, and producer Tom Scholz, the band is a staple of classic rock radio play...

Musical Group, Musical Artist, Composer

Popular Apps based on Freebase

- × Google "Explore More"
- × [Google Refine](#)
- × [Powerset](#) - a semantic search engine (purchased by Microsoft)
- × [Freebase Schema Explorer](#)
- × [FMDb](#)
- × http://en.wikipedia.org/wiki/Freebase_%28database%29

The use of Freebase in Google

Google **the office us**

The Office (U.S. TV series) Wikipedia, the free encyclopedia

The Office (U.S. TV series) The Office is an American comedy television series broadcast by the National Broadcasting Company (NBC). It is an adaptation of the previous BBC series of the same name.

The Office (U.S. season 5) Episodes Characters The Office (U.S. season 5)

The Office (TV Series 2005 -) IMDb

The Office (NBC) Wikipedia, the free encyclopedia

The Office (U.S. in Times & Air Dates Guide) Wikipedia, the free encyclopedia

The Office (TV.com) Wikipedia, the free encyclopedia

The Office (U.S.) Series info

The Office

U.S. TV series

The Office is an American network television series broadcast by the National Broadcasting Company (NBC) as an adaptation of the previous BBC series of the same name. It is a...

First episode: March 24, 2005

Theme song: The Office Theme Song

Writers: Mitch Cullin, Steve Conrad, B. J. Novak, Greg Daniels, Paul Lieberstein, Moe Szilard

Networks: NBC, Official Network Network

Program creator: Ricky Gervais, Stephen Merchant, Greg Daniels

Cast

Steve Conrad, Mitch Cullin, Greg Daniels, Paul Lieberstein, Moe Szilard, Ricky Gervais, Stephen Merchant, Greg Daniels

People also search for

THE OFFICE

Database Project - Goals

- × Project goal: to tackle and resolve **real-life** DB related development issues
- × So what do we need to do:
 - + Design database
 - + Load data / Support updates
 - + Think of an application
 - + Build application
 - + Test

Database Project - Goals (2)

- × What to focus on:
 - + Database
 - + Data Populating/ Updating
 - + Usability
 - + **WOW factor**

Database Project - Features

- ✦ Think your self! Any idea is acceptable
- ✦ Requirements:
 - + Search for specific entities (artist, player..)
 - + Add / Edit / Remove data manually (not just massive import)
 - + Support "Freebase" import: (what happens if you import a new file? You should be able to "skip" existing info..)
 - + **Interesting application**

Database project - DATA????

- ✦ Freebase Data dumps
http://wiki.freebase.com/wiki/Data_dumps
- ✦ Quad Dump ("low level" dump)
<source>, <property>, <destination>, <value>
- ✦ **Simple Topic Dump**
for instance, the "beer.tsv" header:
name, id, beer_style, first_brewed, alcohol_content.....



Database project - DATA????

- ✦ There are 3 files. You do not need to use all..
 - freebase-datadump-tsv.tar.bz2
 - freebase-datadump-quadruples.tsv.bz2
 - freebase-simple-topic-dump.tsv.bz2 (each is a different format..)



- ✦ Additional stuff can be fetched (offline/online) via "Freebase API"
<http://wiki.freebase.com/wiki/Developers>
http://wiki.freebase.com/wiki/Google_API_Client_Libraries

Database project - DATA????

- ✦ We will focus on data from the **TV domain in freebase**
 - ✦ A subset
- ✦ And **related data from another origin**
 - ✦ Other Freebase topics (preferably weird ones)
 - ✦ Other freely available online databases
- ✦ Along with **manual data**



Example

- ✦ Assume we work with the sports domain
- ✦ Create an online (legal..) gambling app
- ✦ You should add/remove/edit all players, sports, games....
- ✦ You should add your own data such odds, bets...
- ✦ Link the data with data from some sports website

17

Past years projects

The screenshot shows the SODA Music Store Manager interface. It includes a search bar, a search results table, and a 'Show Song List' section.

Album ID	Album Name	Artist	Year	Genre	Length
123456	album name 01	led zeppl...	01	rock & roll	76:44
911119	album and one	led zeppl...	0	rock & roll	56:04
1264973	white letter 6s	led zeppl...	1965	classic ro...	75:00
1151298	the one month...	led zeppl...	1973	rock & roll	67:11

Below the table, there is a 'Show Song List' section with a table:

Track	Song name	Artist	Length
1	great baby	Victorae Artista	4:39
2	I can't get you	Victorae Artista	6:30
3	don't get me off	Victorae Artista	3:59
4	you shok me	Victorae Artista	10:42
5	how many men	Victorae Artista	10:40

18

Past years projects



19

Past years projects

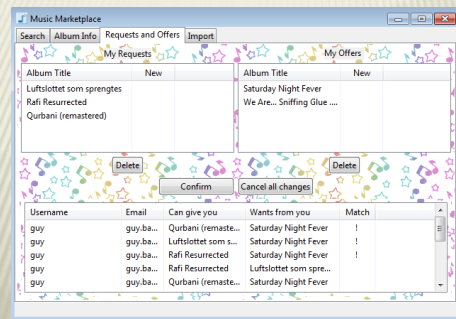


20

Past years projects



Past years projects



22

Tips

- ✦ First:
 - understand the format..
 - understand what you want to do..
- ✦ Database key should always be INTEGER... not a string.... (i.e. you would need to assign it..)
- ✦ Don't forget to support manual edit of ALL data (add/update/remove) – e.g. artists/categories/values...

Database Project - Bureaucracy

- ✦ Hard work, but real.
- ✦ Work in groups of 4
- ✦ Submission database is MySQL in TAU
- ✦ Java, SWT (or Swing/AWT)
- ✦ **Thinking out of the box will be rewarded**

Database Project - Requirements

- ✦ (at least) 150K records table
- ✦ If you use other domains, don't be afraid to import much more data (when its available).