

CC3201-1

BASES DE DATOS

PRIMAVERA 2016

Clase 1: Introducción

Aidan Hogan

aidhog@gmail.com

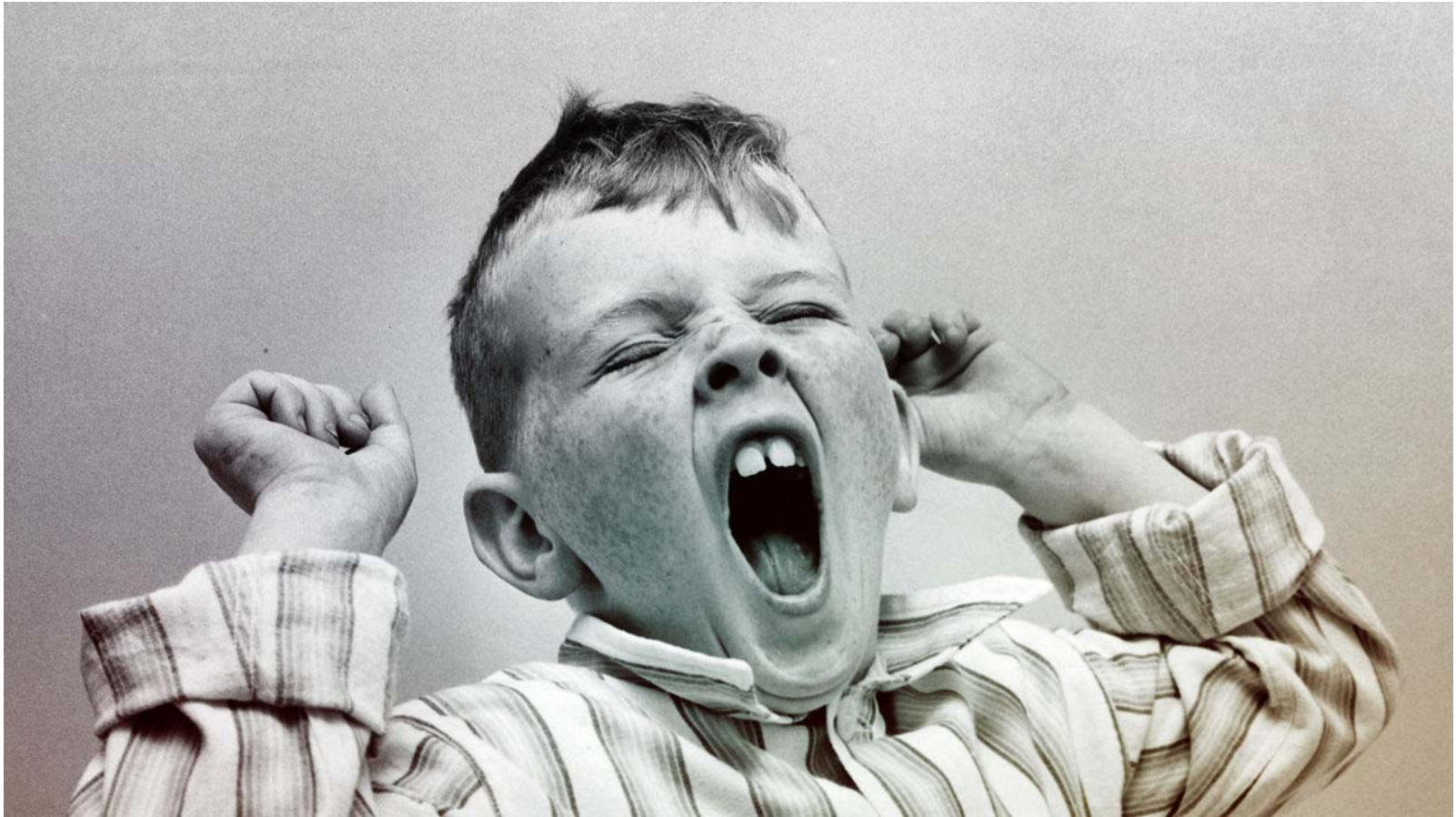
Información General

- Profesor: Aidan Hogan
 - Email: ahogan@dcc.uchile.cl
- Profesor Auxiliar: Sebastián Ferrada
 - Email: scferradaa@gmail.com
- Ayudantes:
 - Ismael Álvarez
 - Hernán Sarmiento

POR QUÉ NECESITAN ESTE CURSO

Un día cualquiera: 09:15

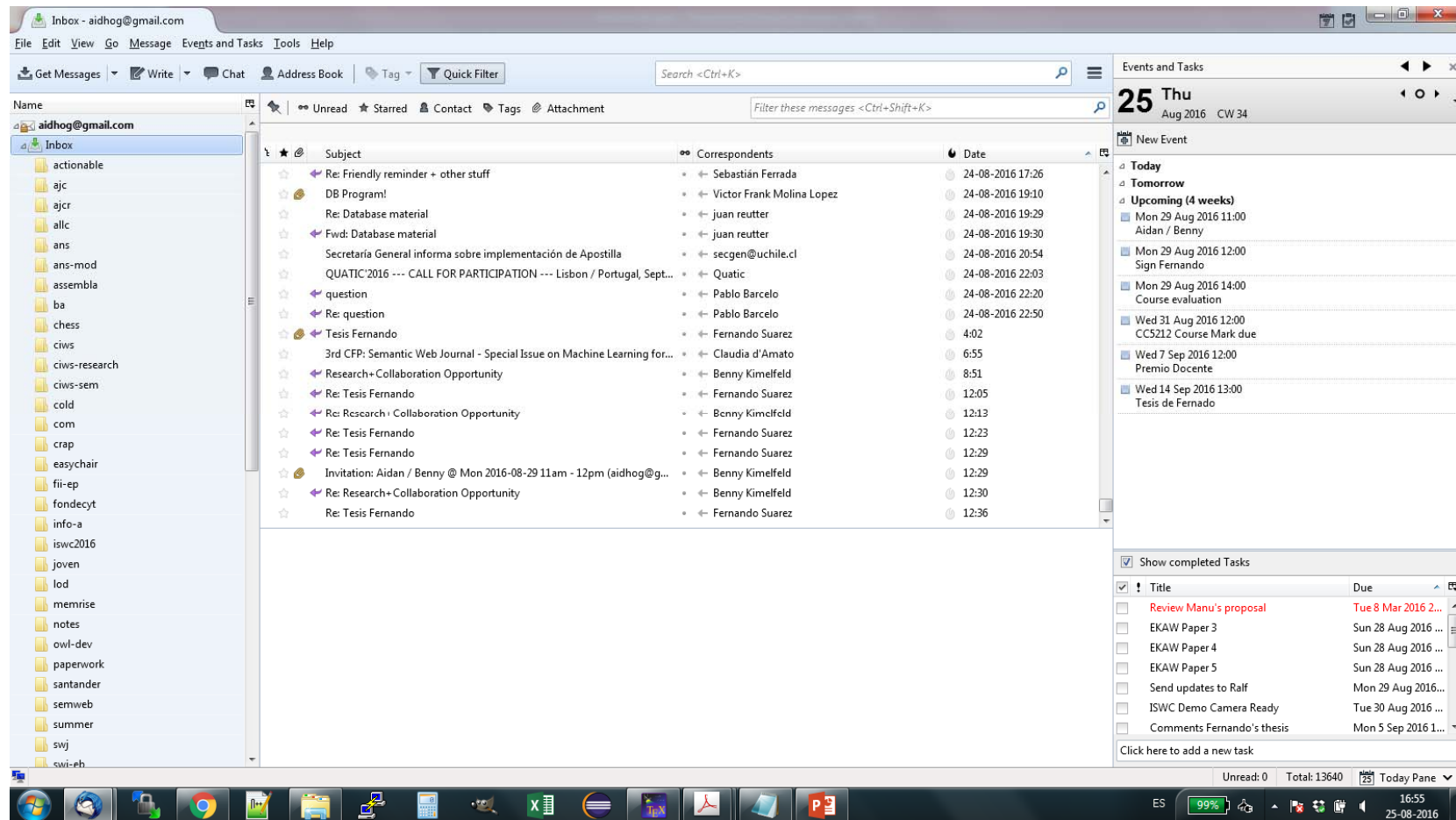
Despierto



(Bostezo.)

Un día cualquiera: 09:35

Reviso el correo



Nada urgente, ¡uf!

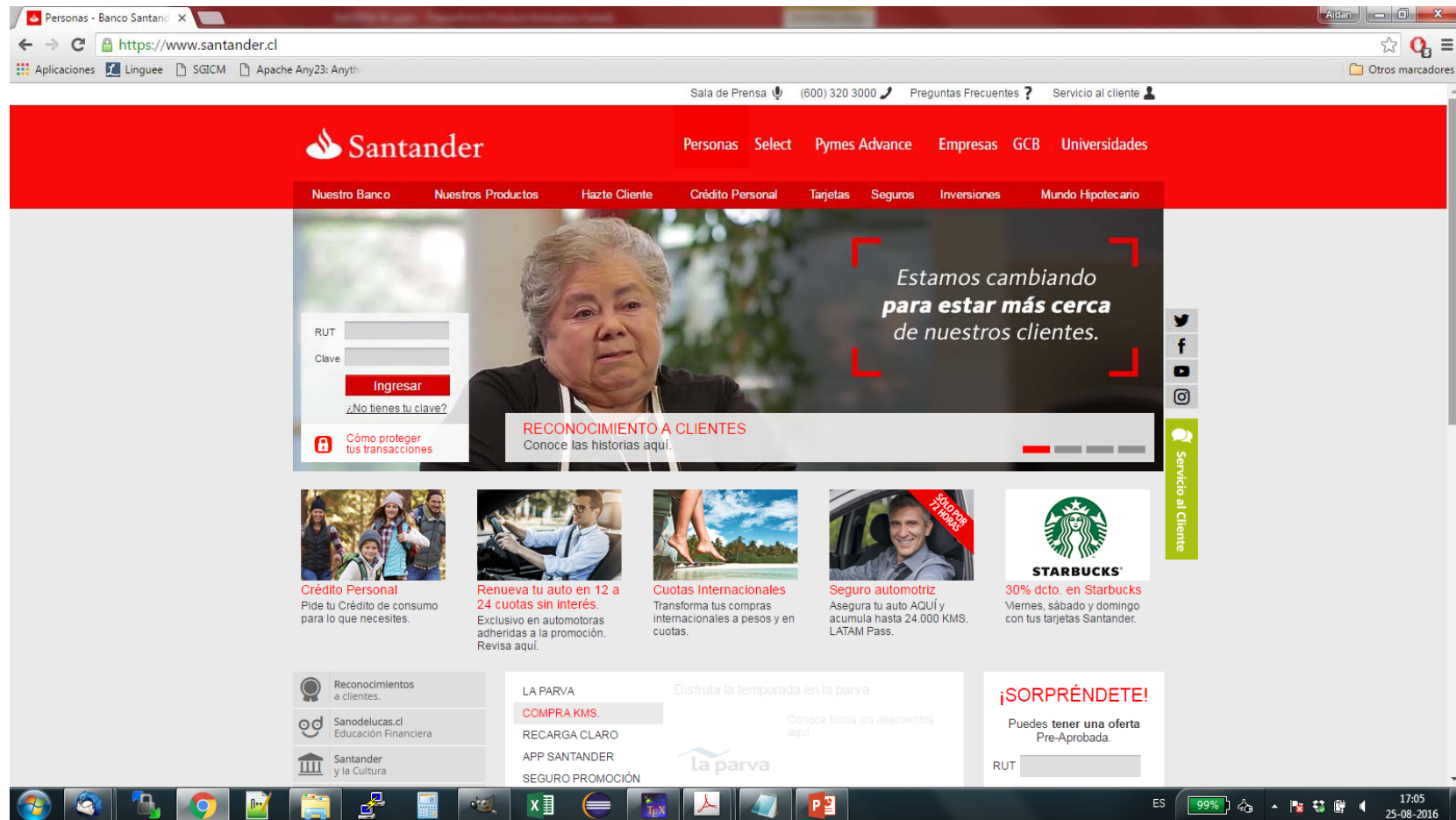
Un día cualquiera: 09:50

Café: pago con tarjeta



Mmm. Café.

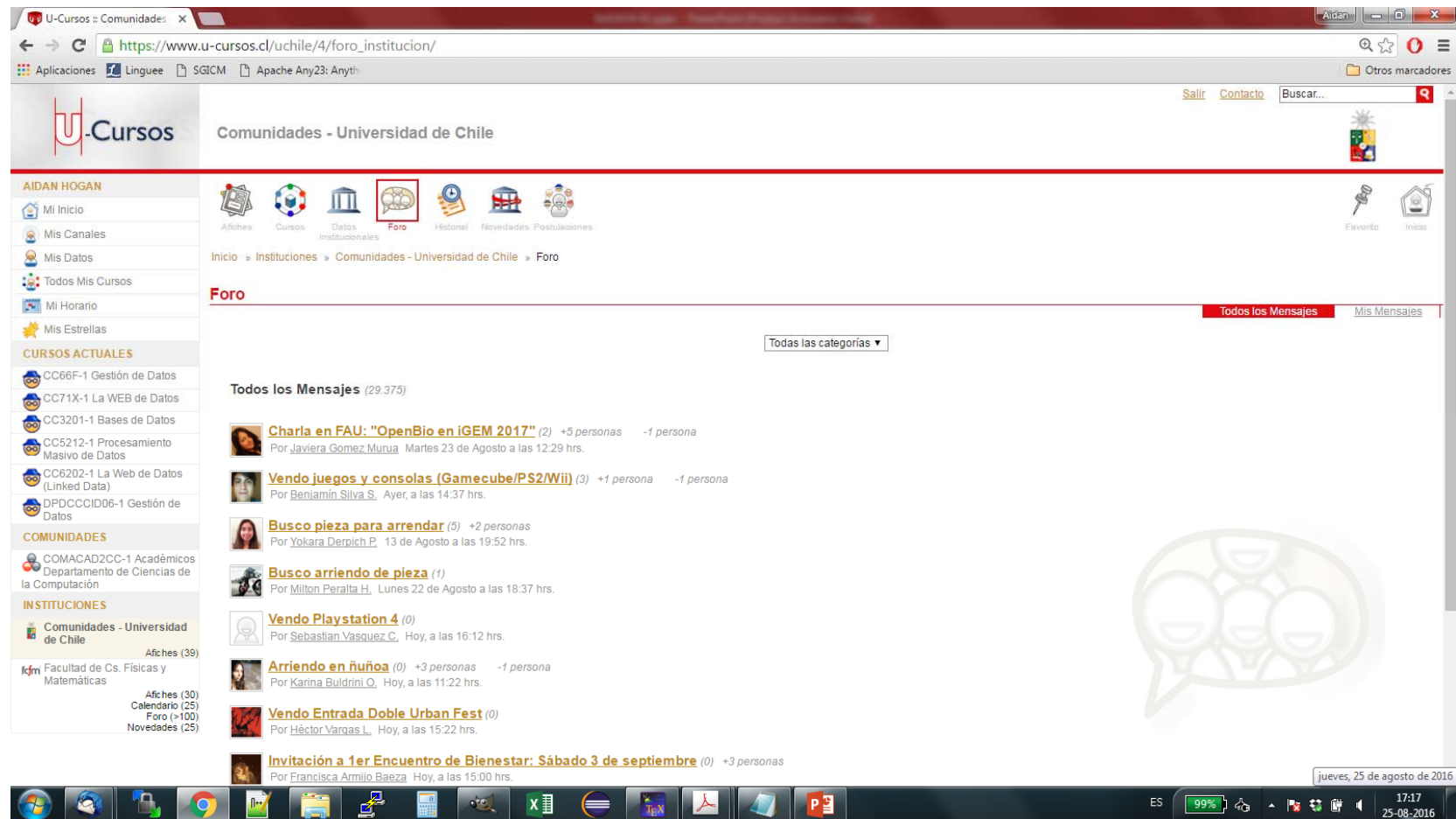
Un día cualquiera: 10:15
Me meto al banco (me pagaron?)



Sí. Me pagaron.

Un día cualquiera: 10:20

Reviso u-cursos (algo en el foro?)



The screenshot shows a web browser window with the URL https://www.u-cursos.cl/uchile/4/foro_institucion/. The page is titled "Comunidades - Universidad de Chile" and features a sidebar with navigation links for "Mi Inicio", "Mis Canales", "Mis Datos", "Todos Mis Cursos", "Mi Horario", and "Mis Estrellas". The main content area is labeled "Foro" and displays a list of messages under the heading "Todos los Mensajes (29.375)". The messages include:

- Charla en FAU: "OpenBio en iGEM 2017"** (2) +5 personas -1 persona. Por [Javiera Gomez Murua](#). Martes 23 de Agosto a las 12:29 hrs.
- Vendo juegos y consolas (Gamecube/PS2/Wii)** (3) +1 persona -1 persona. Por [Benjamin Silva S.](#) Ayer, a las 14:37 hrs.
- Busco pieza para arrendar** (5) +2 personas. Por [Yokara Derpich P.](#) 13 de Agosto a las 19:52 hrs.
- Busco arriendo de pieza** (1). Por [Milton Peralta H.](#) Lunes 22 de Agosto a las 18:37 hrs.
- Vendo Playstation 4** (0). Por [Sebastian Vasquez C.](#) Hoy, a las 16:12 hrs.
- Arriendo en ñuñoa** (0) +3 personas -1 persona. Por [Karina Buldrini O.](#) Hoy, a las 11:22 hrs.
- Vendo Entrada Doble Urban Fest** (0). Por [Héctor Vargas L.](#) Hoy, a las 15:22 hrs.
- Invitación a 1er Encuentro de Bienestar: Sábado 3 de septiembre** (0) +3 personas. Por [Francisca Armijo Baeza](#) Hoy, a las 15:00 hrs.

The bottom of the screen shows a Windows taskbar with various application icons and a system clock indicating the date as Thursday, August 25, 2016, at 17:17.

No, salvo ...

Un día cualquiera: 10:30

IMDb (Mr. Robot ... ¿es bueno?)

The screenshot shows the IMDb website for the TV series "Mr. Robot". The page features a search bar at the top, navigation tabs for Movies, TV & Showtimes, Celebs, Events & Photos, News & Community, and Watchlist. The main content area displays the series title "Mr. Robot" with a rating of 8.8/10 and a "Rate This" button. Below the title is an "Episode Guide" section with 23 episodes. A large video player shows a trailer for "MR. ROBOT season 2.0". To the right, a "Top-Rated Episodes" section lists three episodes: S2.E7 "eps2.5_h4ndshake.sme" (9.6), S1.E8 "eps1.7_wh1ter0se.m4v" (9.5), and S1.E1 "eps1.0_hellofriend.mov" (9.4). The bottom of the page shows a taskbar with various application icons and a system tray with a battery level of 99% and the date 25-08-2016.

Mr. Robot (TV Series 2015) X

www.imdb.com/title/tt4158110/

Aplicaciones | Linguee | SGICM | Apache Any23: Anyth

Otros marcadores

IMDb

Find Movies, TV shows, Celebrities and more...

All

IMDb Pro | Help | f | t | i

Sign in with Facebook | Other Sign in options

Movies, TV & Showtimes | Celebs, Events & Photos | News & Community | Watchlist

FULL CAST AND CREW | TRIVIA | USER REVIEWS | IMDbPro | MORE | SHARE

Mr. Robot ★ 8.8 166,498 Rate This

TV-MA | 49min | Crime, Drama, Thriller | TV Series (2015-)

Episode Guide
23 episodes

CONTROL IS AN ILLUSION

MR. ROBOT
season 2.0 7.13

1:31 Trailer

8 VIDEOS | 360 IMAGES

Mr. Robot follows Elliot, a young programmer who works as a cyber-security engineer by day and a vigilante hacker by night. Elliot finds himself at a crossroad when the mysterious leader of... [See full summary >](#)

Creator: Sam Esmail

Stars: Rami Malek, Christian Slater, Portia Doubleday | [See full cast & crew >](#)

Reviews
350 user | 58 critic

Popularity
5 (1)

Episodes

Top-Rated Episodes

S2.E7
eps2.5_h4ndshake.sme
Mr. Robot and Elliot try to keep the peace. Fsociety releases a video honoring Uncle Sam. Darlene acts upon an old desire.
★ 9.6 Rate this

S1.E8
eps1.7_wh1ter0se.m4v
Allsafe is hacked and Elliot is meeting White Rose. Tyrell is interrogated by the police and Elliot remembers who he really is.
★ 9.5 Rate this

S1.E1
eps1.0_hellofriend.mov
A notorious hacker takes an interest in cyber-security engineer and vigilante-styled computer hacker Elliot; and an evil corporation is hacked.
★ 9.4 Rate this

[See more top-rated episodes >](#)

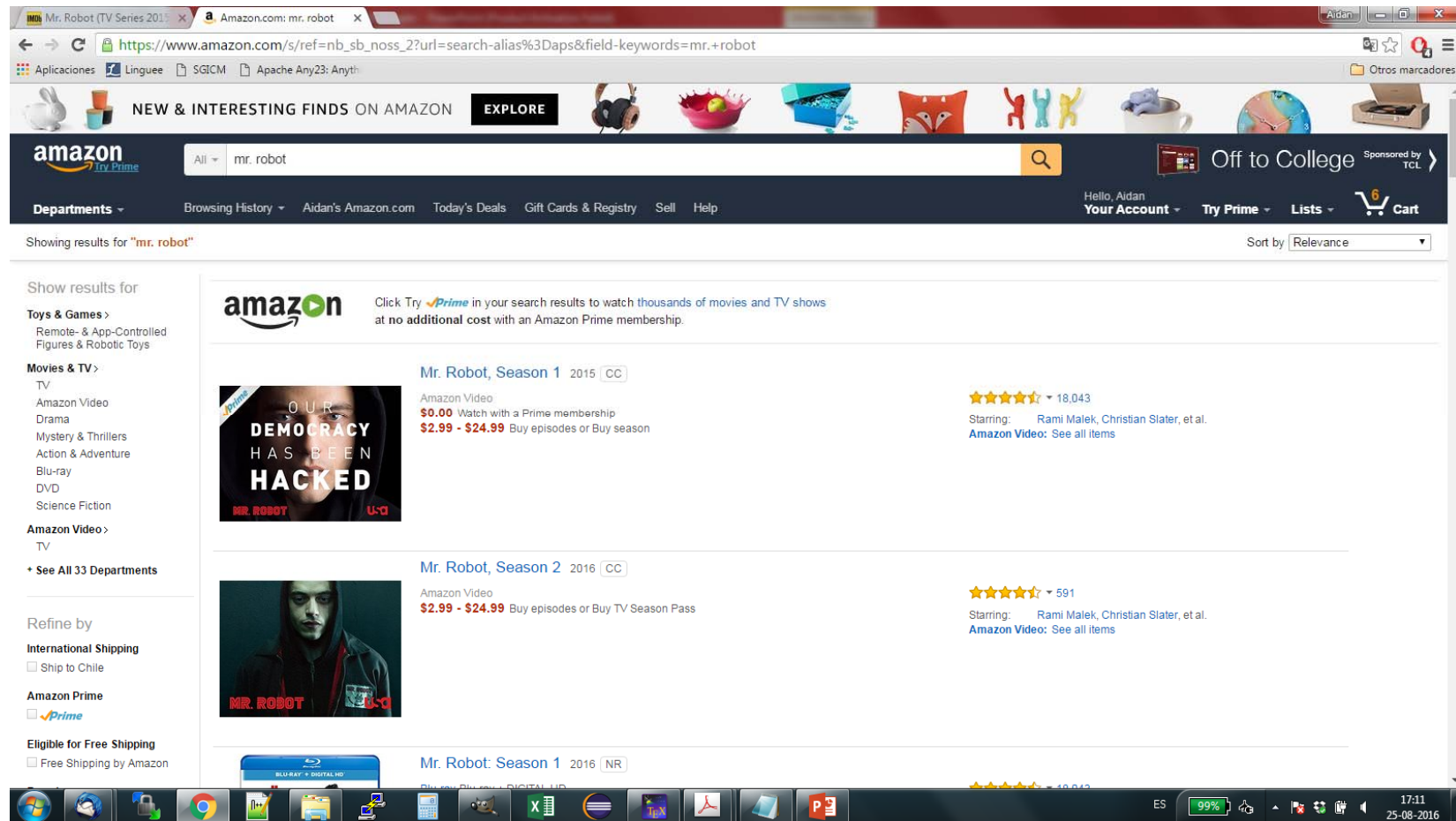
Iconic Back-to-School Movies

ES 99% 17:09 25-08-2016

Sí.

Un día cualquiera: 10:35

Amazon (Mr. Robot ... ¿cuánto cuesta?)



Demasiado.

Un día cualquiera: 10:36

ThePirateBay (Mr. Robot ... me pagaron pero ...)

The screenshot shows a web browser window with the URL <https://thepiratebay.org/search/mr%20robot/0/99/0>. The search results are displayed for the query "mr robot". The results table shows various torrents, including "Mr Robot S02E07 2016 HDTV x264 - iFT", "Mr Robot S02E08 2016 HDTV x264 - iFT", "Mr Robot S02E08 2016 HDTV x264 - FUM", "Mr Robot S02E07 2016 HDTV x264 - FUM", "Mr.Robot.S02E08.HDTV.x264-FUM[ettv]", "Mr Robot S02E07 2016 HDTV x264 - KILLERS", "Mr Robot S02E08 2016 HDTV x264 - KILLERS", "Mr.Robot.S02E07.720p.HDTV.x264-FLEET[PRIME]", "Mr.Robot.S02E06.PROPER.HDTV.x264-KILLERS[ettv]", "Mr.Robot.S02E05.HDTV.x264-FUM[ettv]", "Mr.Robot.S02E03.HDTV.x264-KILLERS[ettv]", "Mr.Robot.S02E02.HDTV.x264-KILLERS[ettv]", "Mr.Robot.S02E04.720p.HDTV.x264-AVS[PRIME]", "Mr.Robot.S02E01.720p.WEBRip.AAC2.0.H.264-KNITTING[ettv]", and "Mr.Robot.S02E08.INTERNAL.720p.HDTV.x264-KILLERS[ettv]". The table includes columns for Type, Name, View, SE, and LE.

Type	Name (Order by: Uploaded, Size, ULed by, SE, LE)	View: Single / Double	SE	LE
Video (HD - TV shows)	Mr Robot S02E07 2016 HDTV x264 - iFT Uploaded Today 19:25, Size 302.03 MiB, ULed by dNLjacka		11905	630
Video (HD - TV shows)	Mr Robot S02E08 2016 HDTV x264 - iFT Uploaded Today 19:25, Size 302.03 MiB, ULed by dNLjacka		10327	669
Video (HD - TV shows)	Mr Robot S02E08 2016 HDTV x264 - FUM Uploaded Today 16:50, Size 302.03 MiB, ULed by jtangerine_pig		10266	977
Video (HD - TV shows)	Mr Robot S02E07 2016 HDTV x264 - FUM Uploaded Today 16:50, Size 302.03 MiB, ULed by jtangerine_pig		9066	635
Video (TV shows)	Mr.Robot.S02E08.HDTV.x264-FUM[ettv] Uploaded Today 06:07, Size 224.34 MiB, ULed by ettv		7551	1776
Video (TV shows)	Mr Robot S02E07 2016 HDTV x264 - KILLERS Uploaded Today 16:21, Size 305.08 MiB, ULed by ndrogo111		7475	645
Video (TV shows)	Mr Robot S02E08 2016 HDTV x264 - KILLERS Uploaded Today 16:21, Size 305.08 MiB, ULed by ndrogo111		7060	725
Video (TV shows)	Mr.Robot.S02E07.720p.HDTV.x264-FLEET[PRIME] Uploaded 08-18 05:09, Size 572.38 MiB, ULed by Simmoncelli		4300	465
Video (TV shows)	Mr.Robot.S02E06.PROPER.HDTV.x264-KILLERS[ettv] Uploaded 08-11 05:49, Size 267.69 MiB, ULed by ettv		4257	402
Video (TV shows)	Mr.Robot.S02E05.HDTV.x264-FUM[ettv] Uploaded 08-04 06:25, Size 246.69 MiB, ULed by ettv		3906	197
Video (TV shows)	Mr.Robot.S02E03.HDTV.x264-KILLERS[ettv] Uploaded 07-21 05:21, Size 245.01 MiB, ULed by ettv		3783	274
Video (TV shows)	Mr.Robot.S02E02.HDTV.x264-KILLERS[ettv] Uploaded 07-14 06:05, Size 174.19 MiB, ULed by ettv		3354	253
Video (TV shows)	Mr.Robot.S02E04.720p.HDTV.x264-AVS[PRIME] Uploaded 07-28 05:30, Size 830.42 MiB, ULed by Simmoncelli		3262	313
Video (HD - TV shows)	Mr.Robot.S02E01.720p.WEBRip.AAC2.0.H.264-KNITTING[ettv] Uploaded 07-11 05:16, Size 957.96 MiB, ULed by EtHD		2945	97
Video (HD - TV shows)	Mr.Robot.S02E08.INTERNAL.720p.HDTV.x264-KILLERS[ettv] Uploaded Today 05:33, Size 700.68 MiB, ULed by EtHD		2794	1059

Listo. Pero tengo hambre ...

Un día cualquiera: 10:52

Al supermercado (¿cuánto cuesta?)



Barrato.

Un día cualquiera: 10:55

Al supermercado (esperando en la fila ...)



Meowth? Fomé.

Un día cualquiera: 10:57

Al supermercado (uso mi tarjeta de fidelidad)



Acumulas puntos? **Sí.**

Un día cualquiera: 11:00

Desayuno



Un día cualquiera: antes que las 11:00

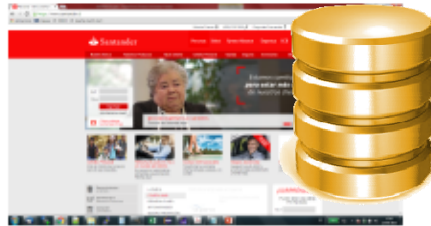
Estas actividades tienen algo en común?

Un día cualquiera: 09:15
Despierto



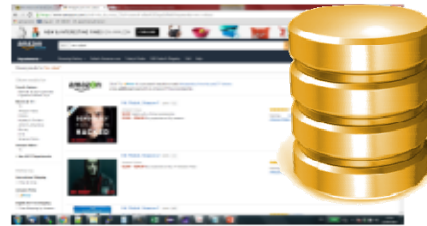
(Bostezo.)

Un día cualquiera: 10:15
Me meto al banco (me pagan?)



Sí. Me pagan.

Un día cualquiera: 10:35
Amazon (Mr. Robot ... ¿cuánto cuesta?)



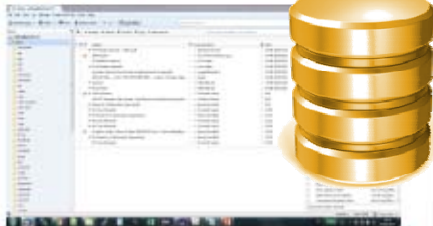
Demasiado.

Un día cualquiera: 10:55
Al supermercado (esperando en la fila ...)



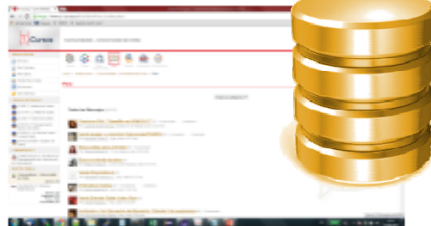
...

Un día cualquiera: 09:35
Reviso el correo



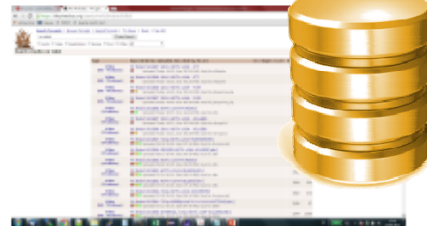
Nada urgente, ¡uf!

Un día cualquiera: 10:20
Reviso u-cursos (algo en el foro?)



No, salvo ...

Un día cualquiera: 10:36
ThePirateBay (Mr. Robot ... cuesta demasiado)



Listo. Pero tengo hambre ...

Un día cualquiera: 10:57
Al supermercado (uso mi tarjeta de fidelidad)



Mucha gente.

Un día cualquiera: 09:50
Café: pago con tarjeta



Mmm. Café.

Un día cualquiera: 10:30
IMDb (Mr. Robot ... ¿es bueno?)



Sí.

Un día cualquiera: 10:52
Al supermercado (¿cuánto cuesta?)



Barrato.

Un día cualquiera: 11:00
Desayuno

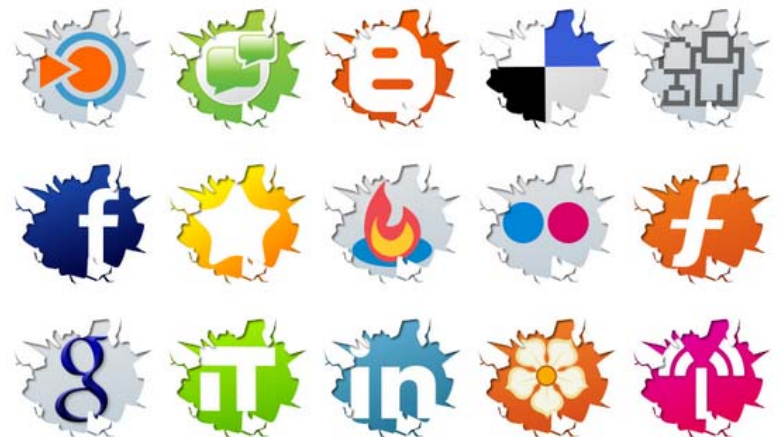


Bases de datos:



Interactuamos con bases de datos
todo el tiempo, todos los días

- Especialmente con la Web:
 - Búsqueda (Google, Bing, Yahoo!, ...)
 - Tiendas (Amazon, eBay, ...)
 - Redes sociales (Facebook, Twitter, ...)
 - Enciclopedias (Wikipedia, IMDb, ...)
 - Bancos
 - Aerolíneas
 - U-cursos
 - ...



Pero frecuentemente solo nos damos cuenta cuando hay un problema

Warning: mysql_connect(): Too many connections in /home/casinome/public_html/forums/dbtech/dbseo/includes/class_db.php on line 2



Database error

The vBulletin database has encountered a problem.

Please try the following:

- Load the page again by clicking the [Refresh](#) button in your web browser.
- Open the www.casinomeister.com home page, then try to open another page.
- Click the [Back](#) button to try another link.

The www.casinomeister.com forum technical staff have been notified of the error, though you may [contact them](#) if the problem persists.

We apologise for any inconvenience.

Database error in vBulletin 2.0.35:

Too many connections

MySQL Error :
Error Number :
Request Date : Tuesday, August 16th 2016 @ 10:21:57 A
Error Date : Tuesday, August 16th 2016 @ 10:21:57 A
Script : http://www.casinomeister.com/forums/on
Referrer :
IP Address : 195.147.229.224
Username : N/A

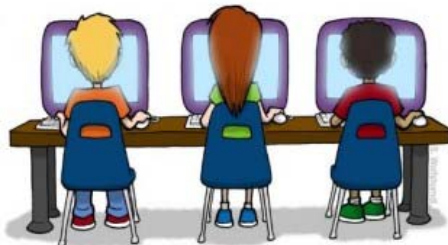
¿Por qué este curso?

Una base de datos es una de nuestras herramientas más importantes



un carpintero : un martillo

un alumno de computación : una base de datos



QUÉ ES UN SISTEMA DE “BASES DE DATOS”

Una base de datos?

*Un ejemplo de una
base de datos?*

Una base de datos?

	PersonID	Address	Address	City	City	FirstName	FirstName	Las
1	1	1040 East Street	1040 East Street	Plateau City	Plateau City	Loretta	Loretta	Bov
2	2	154 Baltic Walk	154 Baltic Walk	Excelsior	Excelsior	Evelyn	Evelyn	Ellio
3	3	952 Tennessee Avenue	952 Tennessee Avenue	Embarcadero	Embarcadero	Harold	Harold	Mcc
4	4	780 Fourth Lane	780 Fourth Lane	Tenderloin	Tenderloin	Chad	Chad	Harr
5	5	1079 Beach Way	1079 Beach Way	Cow Hollow	Cow Hollow	Sandra	Sandra	San
6	6	758 North Lane	758 North Lane	North Beach	North Beach	Kathleen	Kathleen	Rho
7	7	978 Eighth Walk	722 Arrow Lane	Miraloma Park	Columbus	Pamela	Pamela	Nea
8	8	247 Fifth Place	247 Fifth Place	Western Addition	Western Addition	Emily	Emily	Dav
9	9	843 States Street	843 States Street	Noe Valley	Noe Valley	Vernon	Vernon	Carl
10	10	749 Washington Street	749 Washington Street	Civic Center	Civic Center	William	William	Gor
11	11	360 Tennessee Place	360 Tennessee Place	Fisherman's Wharf	Fisherman's Wharf	Gladys	Gladys	Law
12	12	14 Oriental Place	14 Oriental Place	Buena Vista	Buena Vista	Margaret	Margaret	Cole
13	13	668 Lower Avenue	668 Lower Avenue	Diamond Heights	Diamond Heights	Kathleen	Kathleen	Guz
14	14	896 Third Street	896 Third Street	Civic Center	Civic Center	Bertha	Bertha	Pow
15	15	229 Kentucky Place	229 Kentucky Place	Ocean View	Ocean View	Kim	Kim	Gra
16	16	1019 Marvin Gardens Place	1019 Marvin Gardens Place	Potrero Hill	Potrero Hill	Steve	Steve	Cun

Una base de datos?



Una base de datos?

The screenshot displays the IMDb website's layout. At the top is a navigation bar with the IMDb logo, a search bar, and links to 'Movies, TV & Showtimes', 'Celebs, Events & Photos', 'News & Community', and 'Watchlist'. Below this is a section for 'Opening This Week' featuring movie trailers for 'Max Steel', 'Mascots', and 'Shut In'. To the right of these trailers is a list of TV shows including 'Sully', 'When the Bough Breaks', 'The Disappointments Room', 'The Wild Life', 'Author: The JT LeRoy Story', 'Other People', 'Kicks', 'Cameraperson', 'Dancer', and 'Come What May'. Below the trailers is a section titled 'Fall TV: Cliffhanger Catch-Up' with a warning about spoilers and two images of TV show cast members. On the right side, there is a 'Now Playing (Box Office)' section listing movies like 'Don't Breathe', 'Suicide Squad', and 'Pete's Dragon' with their weekend box office earnings. A 'Showtimes' link is also present.

IMDb Find Movies, TV shows, Celebrities and more... All

IMDbPro Help

Sign In with Facebook Other Sign in options

Movies, TV & Showtimes Celebs, Events & Photos News & Community Watchlist

'Max Steel' Official Trailer

'Mascots' Official Trailer

'Shut In' Official Trailer

Browse more trailers »

Opening This Week

- + Sully
- + When the Bough Breaks
- + The Disappointments Room
- + The Wild Life
- + Author: The JT LeRoy Story
- + Other People
- + Kicks
- + Cameraperson
- + Dancer
- + Come What May

See more opening this week »

Now Playing (Box Office)

- + Don't Breathe Weekend: \$19.6M
- + Suicide Squad Weekend: \$12.8M
- + Pete's Dragon Weekend: \$8.6M

Showtimes

Fall TV: Cliffhanger Catch-Up

Here are brief recaps from 10 popular shows just in case you need a reminder on where your favorite shows left off. Warning: *Spoilers ahead!*

Una base de datos?

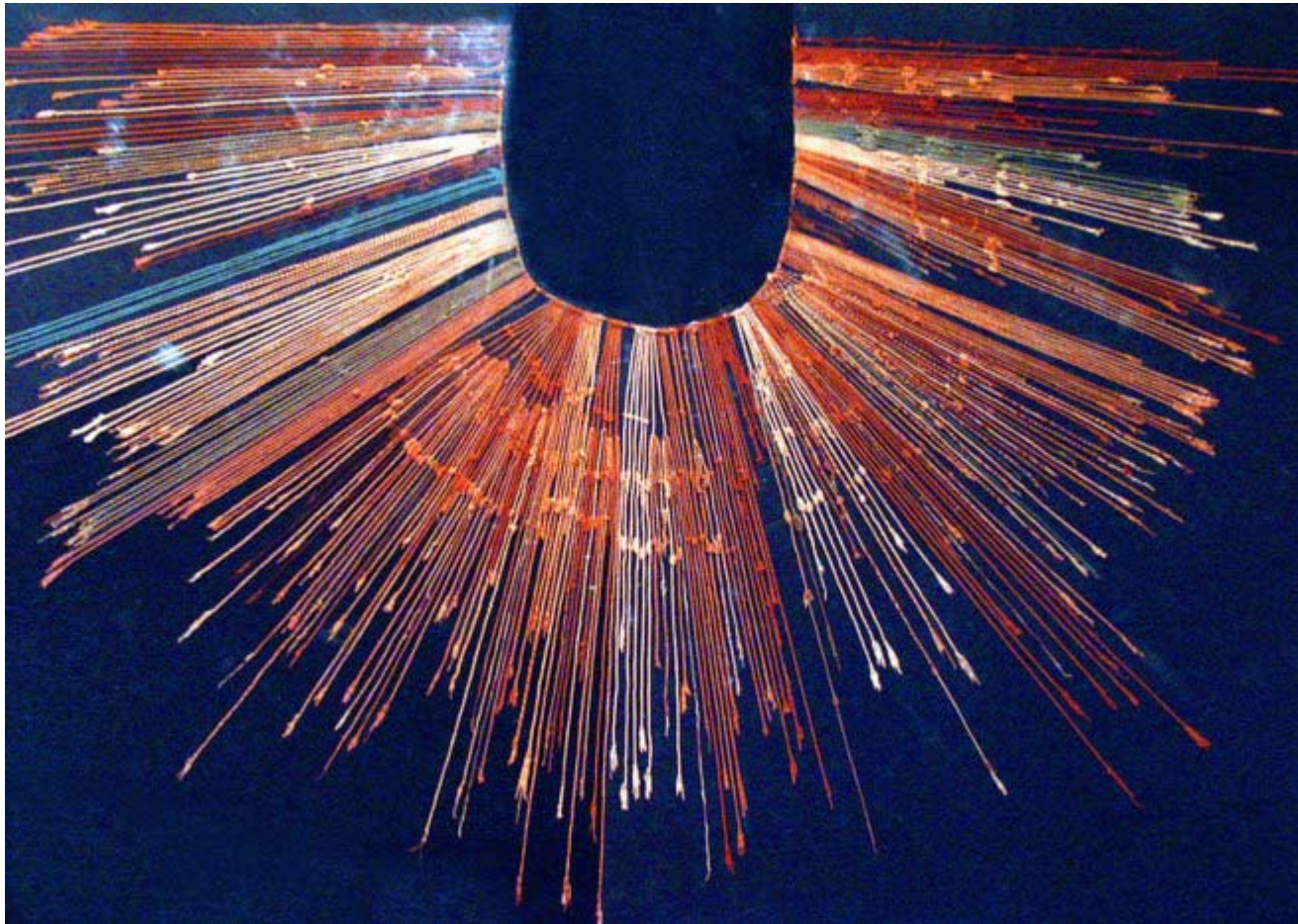
A rectangular search input field with a thin blue border. On the right side of the field is a small microphone icon, indicating voice search functionality.

Google Search

I'm Feeling Lucky

Google.cl offered in: [Español \(Latinoamérica\)](#)

Una base de datos?

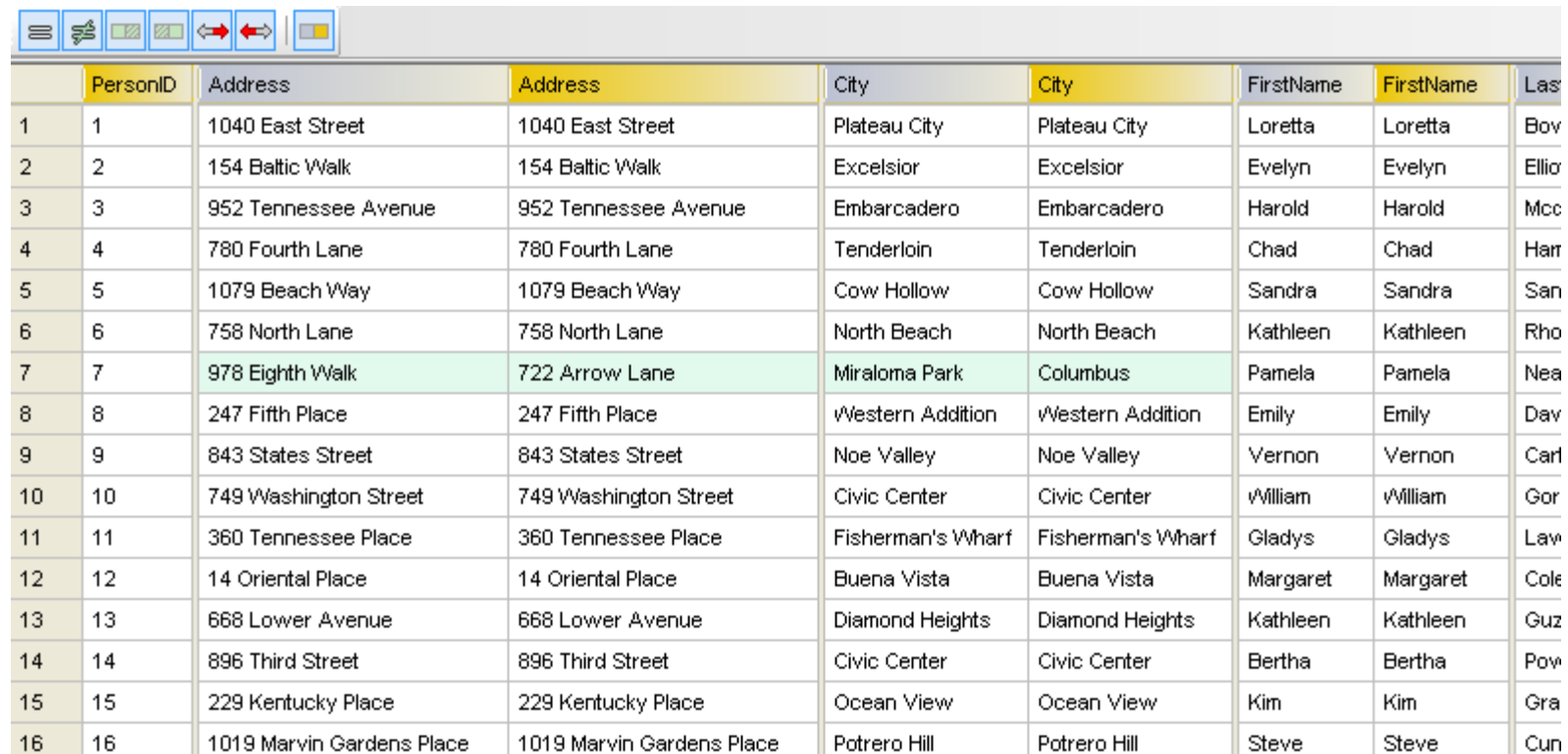


Depende a como se define
una “base de datos”

Aquí, una base de datos es:

- Una colección de datos
(típicamente datos estructurados)
(típicamente datos electrónicos)
organizada en alguna forma
para facilitar **hacer consultas**
en una forma eficiente

Una base de datos?



	PersonID	Address	Address	City	City	FirstName	FirstName	Last
1	1	1040 East Street	1040 East Street	Plateau City	Plateau City	Loretta	Loretta	Bov
2	2	154 Baltic Walk	154 Baltic Walk	Excelsior	Excelsior	Evelyn	Evelyn	Ellio
3	3	952 Tennessee Avenue	952 Tennessee Avenue	Embarcadero	Embarcadero	Harold	Harold	Mcc
4	4	780 Fourth Lane	780 Fourth Lane	Tenderloin	Tenderloin	Chad	Chad	Harr
5	5	1079 Beach Way	1079 Beach Way	Cow Hollow	Cow Hollow	Sandra	Sandra	San
6	6	758 North Lane	758 North Lane	North Beach	North Beach	Kathleen	Kathleen	Rho
7	7	978 Eighth Walk	722 Arrow Lane	Miraloma Park	Columbus	Pamela	Pamela	Nea
8	8	247 Fifth Place	247 Fifth Place	Western Addition	Western Addition	Emily	Emily	Dav
9	9	843 States Street	843 States Street	Noe Valley	Noe Valley	Vernon	Vernon	Carl
10	10	749 Washington Street	749 Washington Street	Civic Center	Civic Center	William	William	Gor
11	11	360 Tennessee Place	360 Tennessee Place	Fisherman's Wharf	Fisherman's Wharf	Gladys	Gladys	Law
12	12	14 Oriental Place	14 Oriental Place	Buena Vista	Buena Vista	Margaret	Margaret	Cole
13	13	668 Lower Avenue	668 Lower Avenue	Diamond Heights	Diamond Heights	Kathleen	Kathleen	Guz
14	14	896 Third Street	896 Third Street	Civic Center	Civic Center	Bertha	Bertha	Pow
15	15	229 Kentucky Place	229 Kentucky Place	Ocean View	Ocean View	Kim	Kim	Gra
16	16	1019 Marvin Gardens Place	1019 Marvin Gardens Place	Potrero Hill	Potrero Hill	Steve	Steve	Cun

(hablando de los datos, no la aplicación ...) Aquí, sí!

Una base de datos?

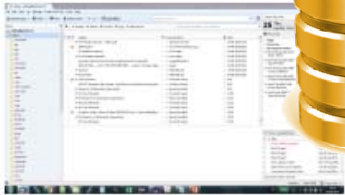


(es un *sistema* de base de datos entonces ...) **Aquí, no!**

Un sistema de bases de datos es:

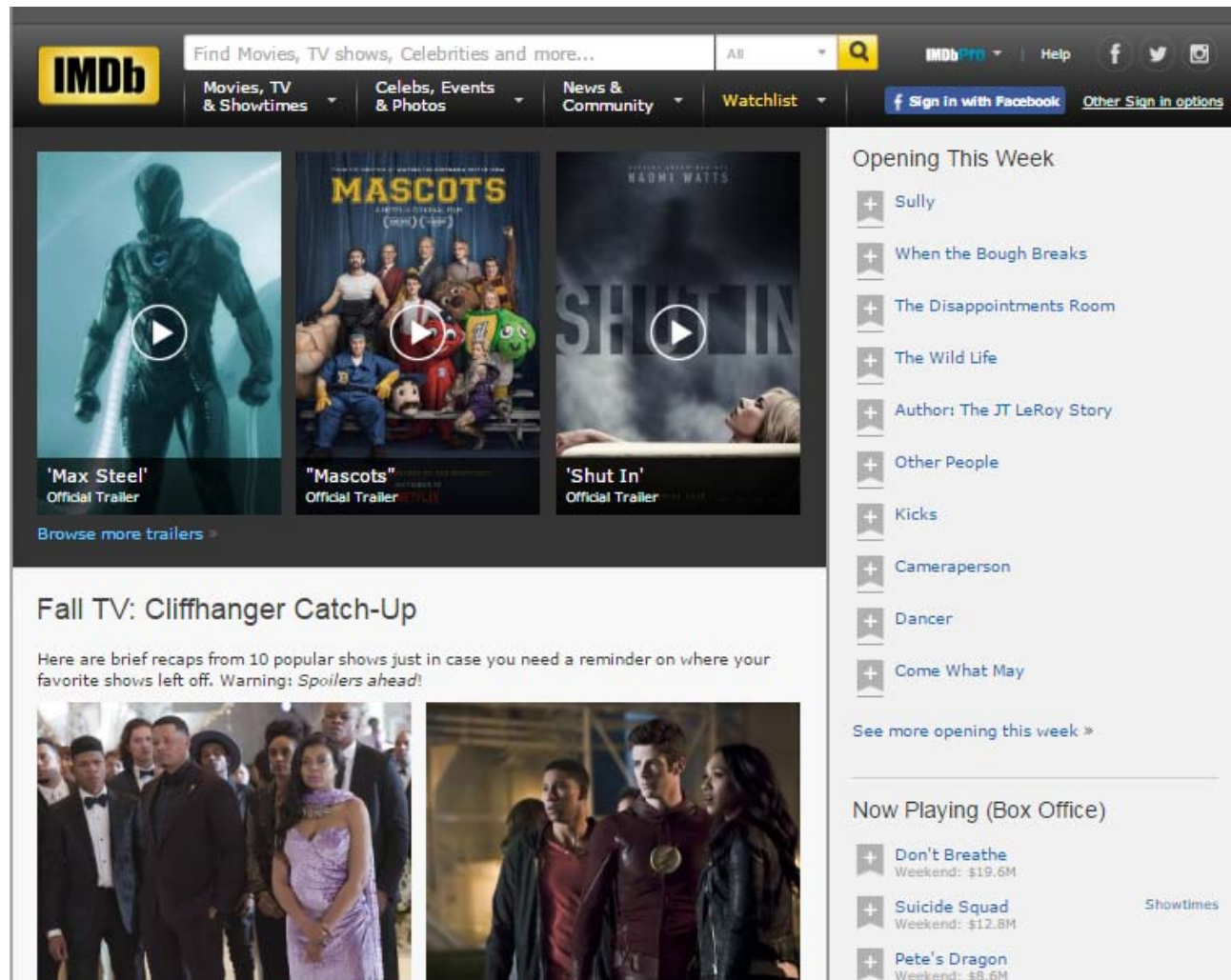
- Un sistema (de software) general
para *manejar*
bases de datos ...
- Facilitar (en una forma general):
 - representar datos,
 - cargar datos,
 - organizar datos,
 - definir datos,
 - actualizar datos,
 - consultar datos,
 - ...
- DBMS: (*DataBase Management System*)

Un sistema general significa que podemos resolver un problema general ...

Un día cualquiera: 09:15 Despierto  (Bostezo.)	Un día cualquiera: 10:15 Me meto al banco (me pagaron?)   Sí. Me pagaron.	Un día cualquiera: 10:35 Amazon (Mr. Robot ... ¿cuánto cuesta?)   Demasiado.	Un día cualquiera: 10:55 Al supermercado (esperando en la fila)   ...
Un día cualquiera: 09:35 Reviso el correo   Nada urgente, ¡uf!	Un día cualquiera: 10:20 Reviso u-cursos (algo en el foro?)   No, salvo ...	Un día cualquiera: 10:36 ThePirateBay (Mr. Robot ... cuesta dema   Listo. Pero tengo hambre ...	Un día cualquiera: 10:57 Al supermercado (uso mi tarjeta de fide   Mucha gente.
Un día cualquiera: 09:50 Café: pago con tarjeta   Mmm. Café.	Un día cualquiera: 10:30 IMDb (Mr. Robot ... ¿es bueno?)   Sí.	Un día cualquiera: 10:52 Al supermercado (¿cuánto cuesta?)   Barrato.	Un día cualquiera: 11:00 Desayuno  

Un sistema de bases de datos. Diez aplicaciones.

Una base de datos?



(es una *aplicación* de base de datos entonces ...) **Aquí, no!**

Una base de datos?



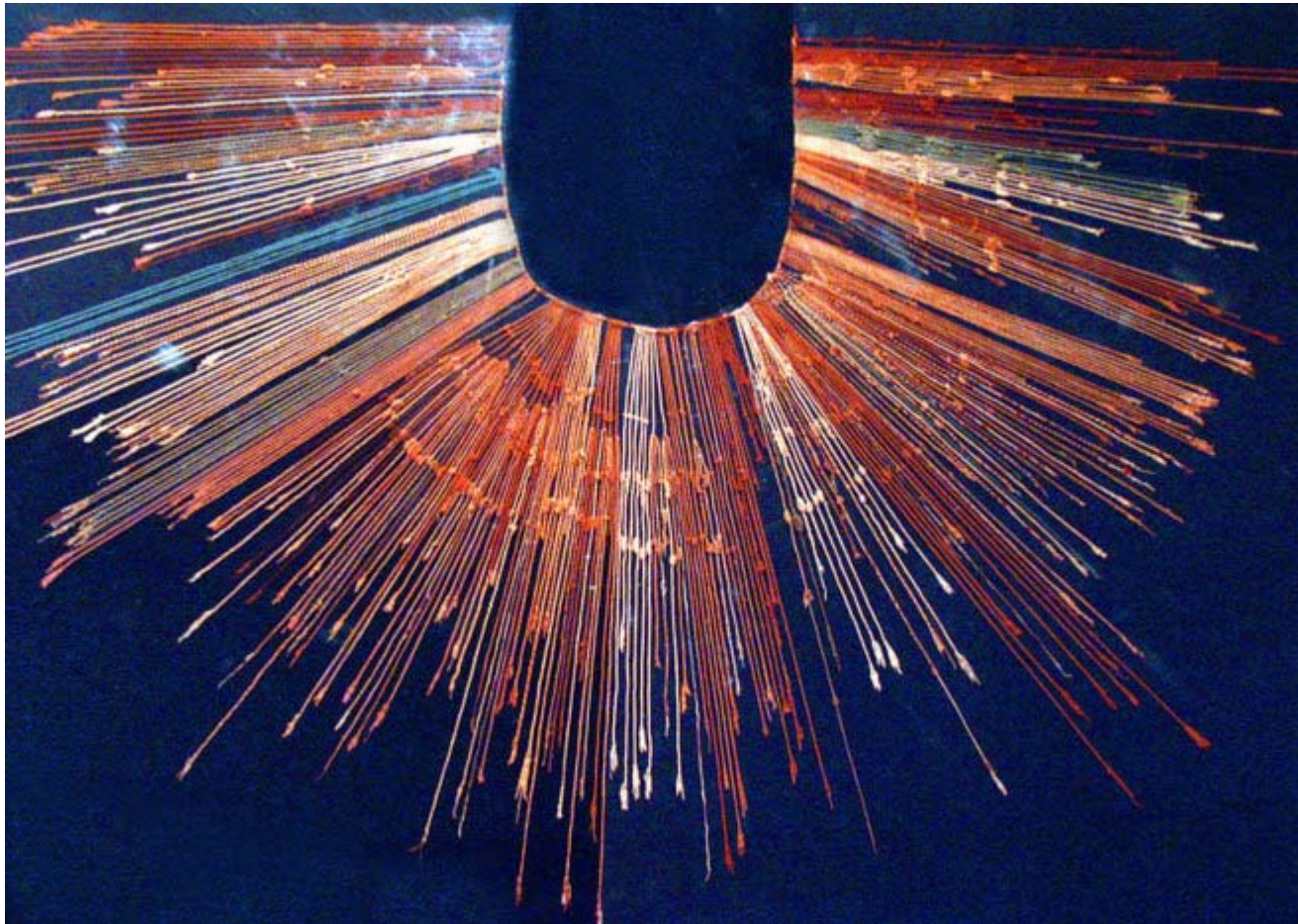
Google Search

I'm Feeling Lucky

Google.cl offered in: [Español \(Latinoamérica\)](#)

(es una *aplicación* de base de datos entonces ...) **Aquí, no!**

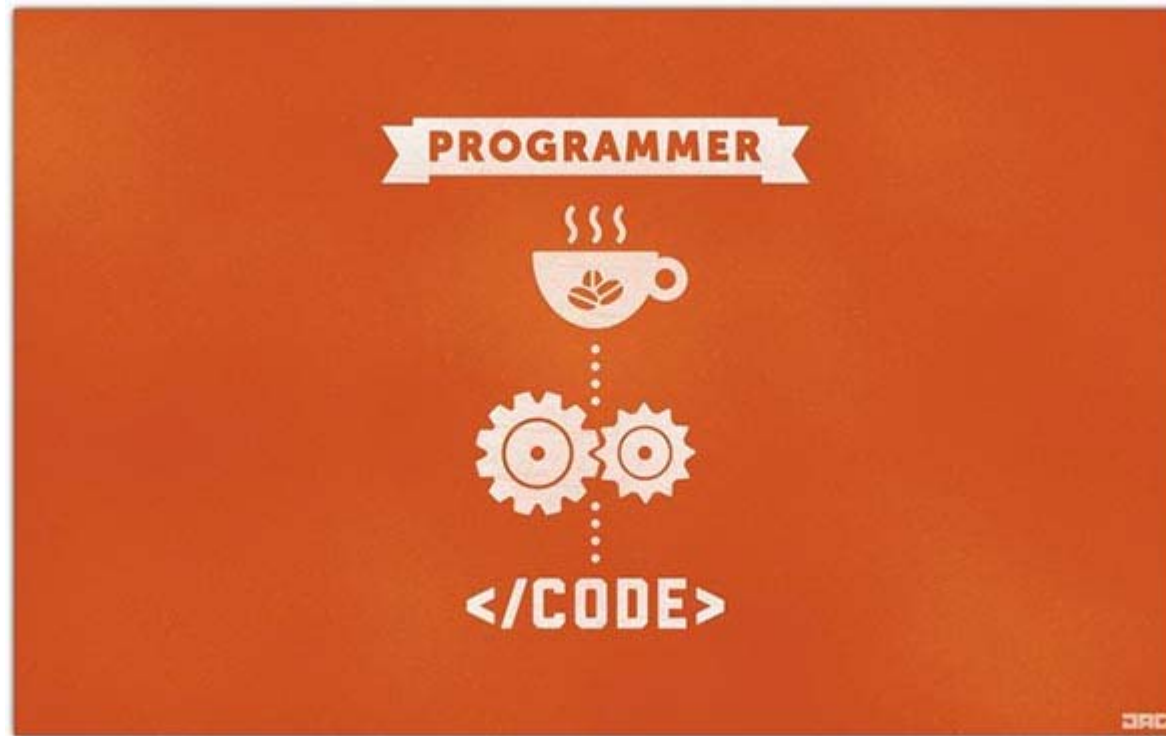
Una base de datos?



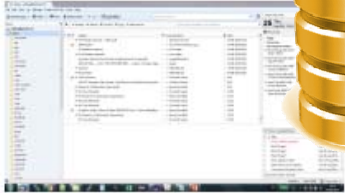
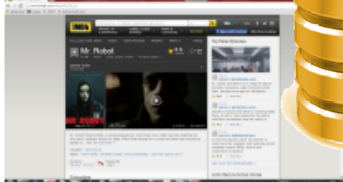
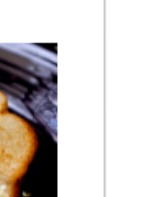
Hmm.

POR QUÉ SE NECESITAN SISTEMAS DE “BASES DE DATOS”

Sé programar en Java, sé programar en Python,
... podría programar algo sin problema!



Intentemos implementar una aplicación sin un sistema de bases de datos (p.e., en Java)

Un día cualquiera: 09:15 Despierto  (Bostezo.)	Un día cualquiera: 10:15 Me meto al banco (me pagaron?)   Sí. Me pagaron.	Un día cualquiera: 10:35 Amazon (Mr. Robot ... ¿cuánto cuesta?)   Demasiado.	Un día cualquiera: 10:55 Al supermercado (esperando en la fila)   ...
Un día cualquiera: 09:35 Reviso el correo   Nada urgente, ¡uf!	Un día cualquiera: 10:20 Reviso u-cursos (algo en el foro?)   No, salvo ...	Un día cualquiera: 10:36 ThePirateBay (Mr. Robot ... cuesta dema   Listo. Pero tengo hambre ...	Un día cualquiera: 10:57 Al supermercado (uso mi tarjeta de fide   Mucha gente.
Un día cualquiera: 09:50 Café: pago con tarjeta   Mmm. Café.	Un día cualquiera: 10:30 IMDb (Mr. Robot ... ¿es bueno?)   Sí.	Un día cualquiera: 10:52 Al supermercado (¿cuánto cuesta?)   Barrato.	Un día cualquiera: 11:00 Desayuno  

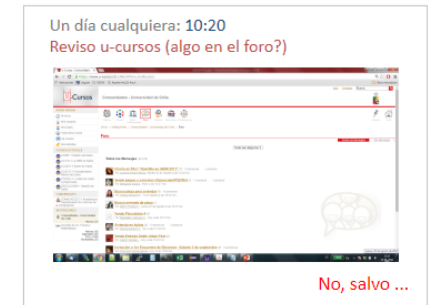
Un sistema de bases de datos. Diez aplicaciones.

Tenemos información de profesores, auxiliares, integrantes y notas parciales en cada curso

profesores.csv		
RUT	Nombre	Curso
...		
24.482.054-9	Aidan Hogan	CC3201
24.482.054-9	Aidan Hogan	CC5212
24.482.054-9	Aidan Hogan	CC6202
...		

alumnos.csv		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	



auxiliares.csv		
RUT	Nombre	Curso
...		
12.412.412-4	Sebastián Ferrada	CC3201
12.412.412-4	Sebastián Ferrada	CC5208
13.123.024-9	Daniel Hernández	CC6202
...		

notas.csv			
RUT	Nombre	Eval	Nota
...			
10.323.634-4	CC3201	Lab1	6.0
10.323.634-4	CC3201	Lab2	4.5
12.323.792-8	CC6202	Examen	3.9
...			

Queremos saber todos los códigos del cursos que toma el alumno “12.323.792-8”

alumnos.csv		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

- En Java, podemos leer todo el archivo, filtrar todas las filas con otros RUTs y entregar solo la información relevante

Algún problema aquí?

Bueno, si los usuarios son impacientes y los archivos grandes ...

<i>Mapa en memoria principal de alumnos.csv</i>	
Clave	↦ Valor
...	↦ ...
10.323.634-4	↦ { (Pia García,CC3201) }
12.323.792-8	↦ { (Juan Ramírez,CC6202), (Juan Ramírez,CC5212) }
...	↦ ...

- En Java, podemos cargar los datos en memoria principal, y utilizar un índice con RUTs como claves (p.e., un “hashmap”)

Algún problema aquí?

Bueno, si los usuarios son impacientes y los archivos no caben en memoria ...

alumnos.csv		
RUT	Nombre	Curso
10.323.634-4	Pia García	CC3201
10.323.634-5	Jorge García	CC3201
...		
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		
...		

Indíce (m. principal)	
RUT	Bloque
10.323.634-4	1
12.323.792-8	2
...	

- En Java, podemos crear bloques de datos ordenados por RUT, y utilizar un índice con el primer RUT en cada bloque

Algún problema aquí?

Bueno, si tenemos que actualizar la tabla con datos nuevos ...

<i>Indíce (m. principal)</i>	
RUT	Bloque
10.323.634-4	1
12.323.792-8	2
...	

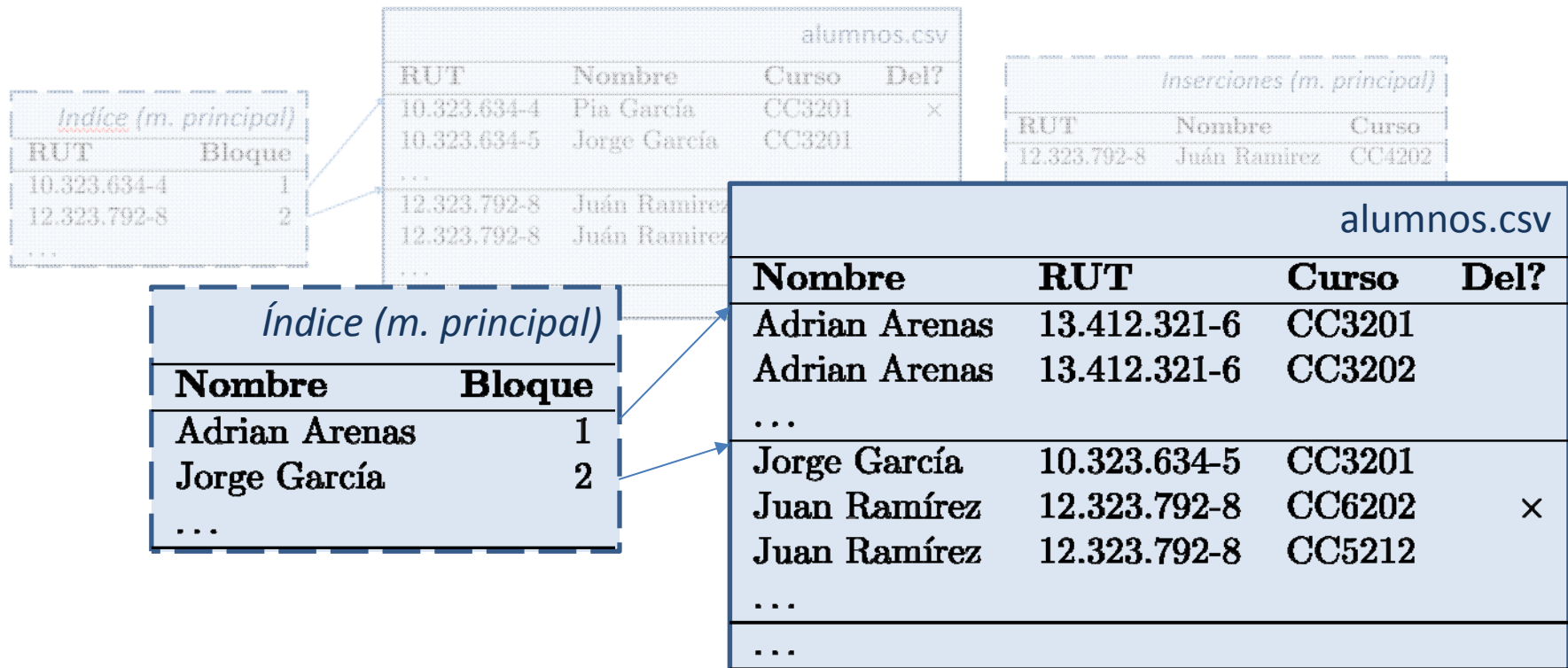
alumnos.csv			
RUT	Nombre	Curso	Del?
10.323.634-4	Pia García	CC3201	×
10.323.634-5	Jorge García	CC3201	
...			
12.323.792-8	Juan Ramírez	CC6202	×
12.323.792-8	Juan Ramírez	CC5212	
...			
...			

<i>Inserciones (m. principal)</i>		
RUT	Nombre	Curso
12.323.792-8	Juan Ramírez	CC4202
...		

- En Java, podemos crear un bloque en memoria principal, o podemos dejar espacio en los bloques para datos nuevos o ...

Algún problema aquí?

Bien, si a veces hay que consultar por el nombre del alumno entonces ...



- En Java, podemos crear otro índice ordenado por nombre ...

Algún problema aquí?

ok ok, si a veces hay que consultar por el nombre del cursos del alumno entonces ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

- En Java, podemos crear otro índice para cursos.csv e implementar “*joins*” entre ambos índices

Algún problema aquí?

... fine, si hay que asegurar que los alumnos solo tienen cursos que aparecen en cursos.csv ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

INSERT alumnos.csv (12.323.792-8,Juan Ramírez,CC3201) ✓
INSERT alumnos.csv (12.323.792-8,Juan Ramírez,BdeDatos) ✗

- En Java, antes que hacer una inserción a alumnos.csv, podemos consultar a cursos.csv para verificar que el curso existe.

Algún problema aquí?

... pues, si hay que permitir quitar cursos ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

DO

```
DELETE alumnos.csv (10.323.634-4,Pia García,CC3201)
...
DELETE cursos.csv (CC3201,Bases de Datos)
```

- En Java, podemos agrupar inserciones y/o borrados para mantener la consistencia de los datos (transacciones)

Algún problema aquí?

... si hay múltiples usuarios actualizando la base de datos al mismo tiempo ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

DO

DELETE	alumnos.csv	(10.323.634-4,Pia García,CC3201)
...		
DELETE	cursos.csv	(CC3201,Bases de Datos)

DO

INSERT	alumnos.csv	(14.234.723-3,Pablo Silva,CC3201)
---------------	-------------	-----------------------------------

- En Java, hay que aislar transacciones para evitar esta tipa de situación (y otras similares)

Algún problema aquí?

... si hay que contar el numero de cursos que cada alumno toma u otra formas de consultas ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

```
SELECT RUT, COUNT(Curso) FROM alumnos GROUP BY RUT
```

- En Java, podemos implementar un lenguaje de consulta general que cubre los rasgos más necesitados

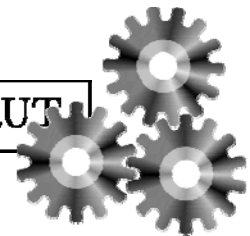
Algún problema aquí?

... si el rendimiento de consultas no basta para los usuarios podemos optimizar ...

alumnos.csv (indexado por RUT y Nombre)		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

cursos.csv (Indexado por Codigo)	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

SELECT RUT, COUNT(Curso) FROM alumnos GROUP BY RUT



- En Java, podemos implementar varias optimizaciones en un planificador de ejecución

Algún problema aquí?

(°□°) 丿 ㄣ ㄣ ㄣ

¡Sí!

- Faltan valores
- Cursos pueden tener más que un nombre
- Tenemos valores como fechas, booleanos, etc., que queremos comparar, ordenar, manipular, sumar ...
- El rendimiento de algunas consultas todavía es terrible
- El cargo de datos todavía es demasiado lento
- No hay suficiente memoria para mantener los índices
- Los administradores quieren agregar columnas nuevas como la carrera de los alumnos
- Los alumnos no deberían tener acceso para cambiar sus notas
- Hay “l33t h4cker\$” que quieren *pwnear* nuestra base de datos para cambiar sus notas
- Tenemos que mantener *backups* en una forma segura

... y si pudiéramos solucionar estos problemas
en una forma general ...



... habríamos implementado
un sistema de bases de datos

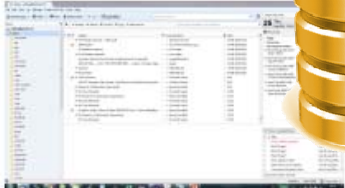
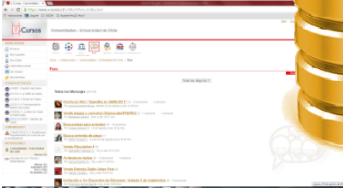
... y si pudiéramos solucionar estos problemas en una forma general ...



(... y probablemente tendríamos una rueda no muy bien reinventada)

... habríamos implementado un sistema de bases de datos

Estos son problemas generales que se encuentran en muchas aplicaciones

Un día cualquiera: 09:15 Despierto  (Bostezo.)	Un día cualquiera: 10:15 Me meto al banco (me pagaron?)   Sí. Me pagaron.	Un día cualquiera: 10:35 Amazon (Mr. Robot ... ¿cuánto cuesta?)   Demasiado.	Un día cualquiera: 10:55 Al supermercado (esperando en la fila)   ...
Un día cualquiera: 09:35 Reviso el correo   Nada urgente, ¡uf!	Un día cualquiera: 10:20 Reviso u-cursos (algo en el foro?)   No, salvo ...	Un día cualquiera: 10:36 ThePirateBay (Mr. Robot ... cuesta dema   Listo. Pero tengo hambre ...	Un día cualquiera: 10:57 Al supermercado (uso mi tarjeta de fide   Mucha gente.
Un día cualquiera: 09:50 Café: pago con tarjeta   Mmm. Café.	Un día cualquiera: 10:30 IMDb (Mr. Robot ... ¿es bueno?)   Sí.	Un día cualquiera: 10:52 Al supermercado (¿cuánto cuesta?)   Barrato.	Un día cualquiera: 11:00 Desayuno  

... muchas aplicaciones importantes

Un día cualquiera: 10:15
Me meto al banco (me pagaron?)



Sí. Me pagaron.



Un sistema de bases de datos es:

- Un sistema (de software) general
para *manejar*
bases de datos ...
- Facilitar (en una forma general):
 - representar datos,
 - cargar datos,
 - organizar datos,
 - definir datos,
 - actualizar datos,
 - consultar datos,
 - ...
- DBMS: (*DataBase Management System*)

Con un DBMS ...

Los **usuarios** se encargan de:

- diseñar la estructura de la base de datos,
- escribir consultas,
- actualizar los datos,
- ...

... **solo las cosas especificas en el contexto de la aplicación especifica.**

Con un DBMS ...

Por debajo, el **DBMS** se encarga de:

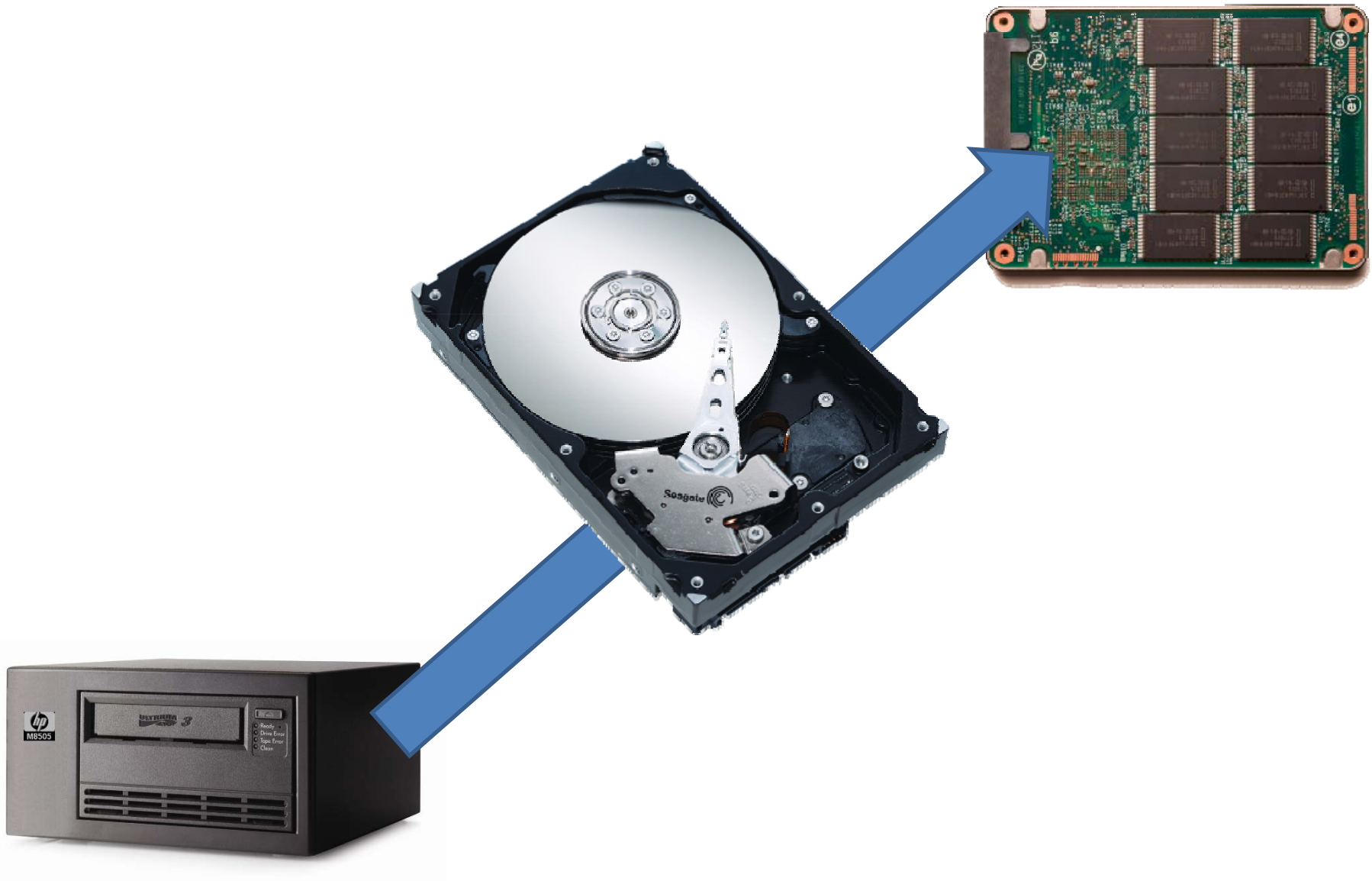
- Almacenaje optimizado
- Indexación
- Motor de consultas
- Optimizar consultas
- Manejo de transacciones
- Manejo de acceso concurrente
- Seguridad
- *¡y mucho más!*

... las cosas generales que se necesitan en el contexto de muchas aplicaciones.

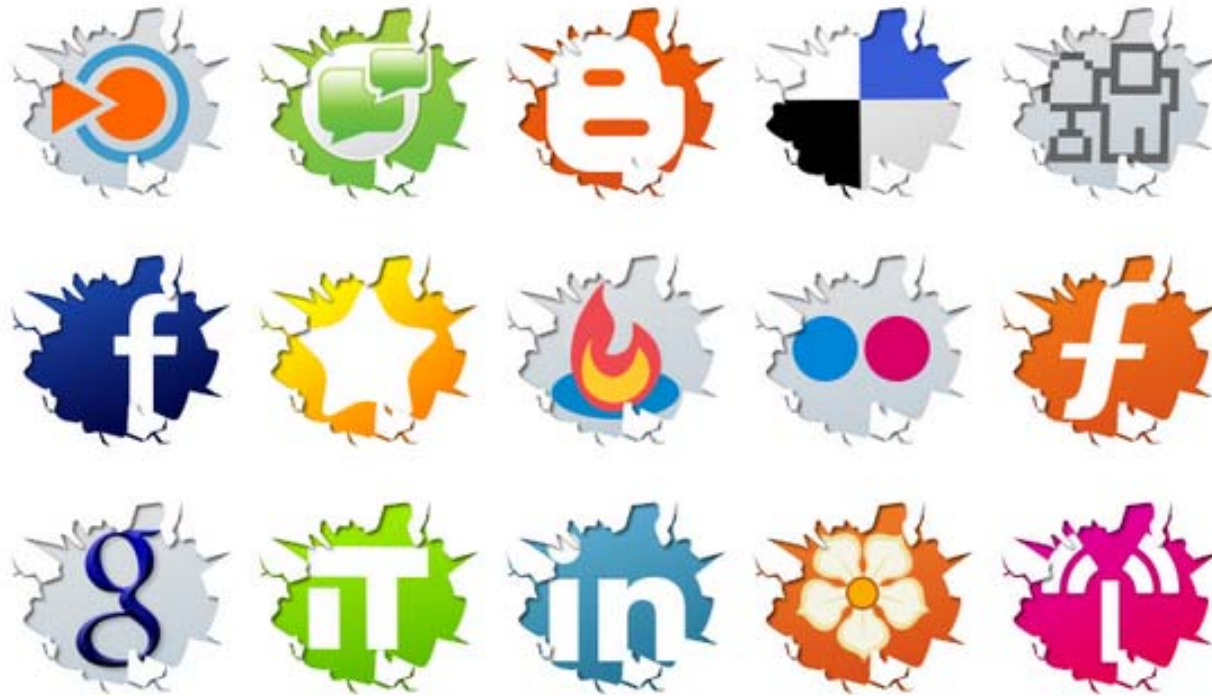
Hay implementaciones con décadas de desarrollo por miles de expertos



Pero DBMS están siempre evolucionando:
tecnología cambia



Pero DBMS están siempre evolucionando:
los requisitos de las aplicaciones cambian



Logo bingo!!

Por ejemplo, distribución es más importante



Diluvio de datos



**BASES DE DATOS SIEMPRE
MODELAN DATOS COMO TABLAS?**

... son siempre modelados así?

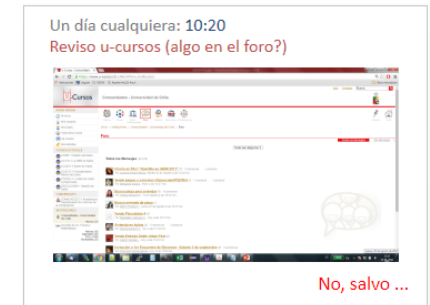
profesores.csv		
RUT	Nombre	Curso
...		
24.482.054-9	Aidan Hogan	CC3201
24.482.054-9	Aidan Hogan	CC5212
24.482.054-9	Aidan Hogan	CC6202
...		

alumnos.csv		
RUT	Nombre	Curso
...		
10.323.634-4	Pia García	CC3201
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

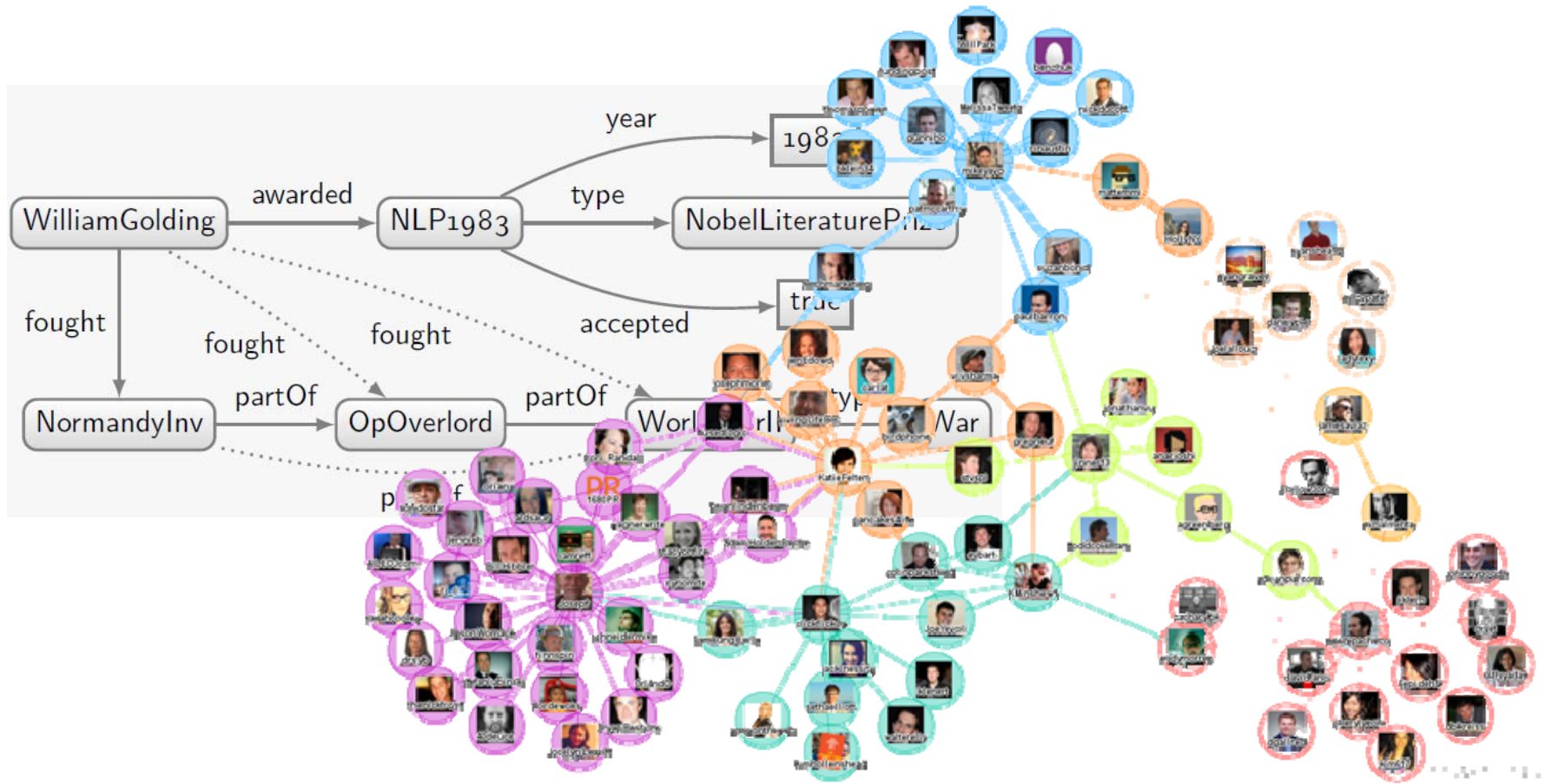
cursos.csv	
Codigo	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

auxiliares.csv		
RUT	Nombre	Curso
...		
12.412.412-4	Sebastián Ferrada	CC3201
12.412.412-4	Sebastián Ferrada	CC5208
13.123.024-9	Daniel Hernández	CC6202
...		

notas.csv			
RUT	Nombre	Eval	Nota
...			
10.323.634-4	CC3201	Lab1	6.0
10.323.634-4	CC3201	Lab2	4.5
12.323.792-8	CC6202	Examen	3.9
...			

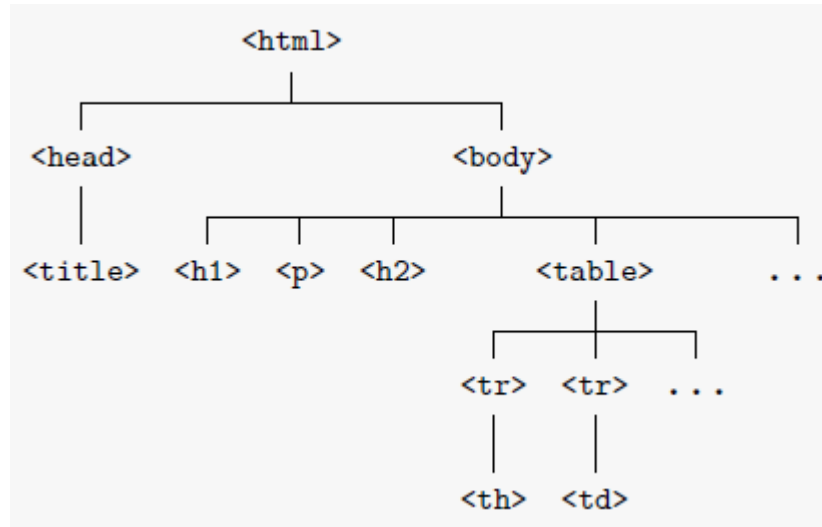


¿Se puede modelar una base de datos como grafos?



¡Sí!

¿Se puede modelar una base de datos como arboles?



JSON

```
{
  "siblings": [
    {"firstName": "Anna", "lastName": "Clayton"},
    {"lastName": "Alex", "lastName": "Clayton"}
  ]
}
```

XML

```
<siblings>
  <sibling>
    <firstName>Anna</firstName>
    <lastName>Clayton</lastName>
  </sibling>
  <sibling>
    <firstName>Alex</firstName>
    <lastName>Clayton</lastName>
  </sibling>
</siblings>
```

¡Sí!

¿Se puede modelar una base de datos como un mapa?

Clave	↦ Valor
...	↦ ...
10.323.634-4	↦ { (Pia García,CC3201) }
12.323.792-8	↦ { (Juán Ramirez,CC6202), (Juán Ramirez,CC5212) }
...	↦ ...

Examples:	Key	Value
Directory	Company Algo-Logic	Phone # (408) 707-3740
Forwarding Tables	IP Address 204.2.34.5	Interface : MAC Address Eth6 : 02:33:29:F2:AB:CC
Data De- duplication	Content Hash XYZ	Storage Block ID 948830038411
Stock Trading	Order ID ATY11217911101	Symbol, Side, Price AAPL, B, 126.75
Graph Search	Virtex v140	Edge List v201, v206, v225

¡Sí!

Bases de Datos Relacional

- Tablas = Un Modelo de Bases de Datos

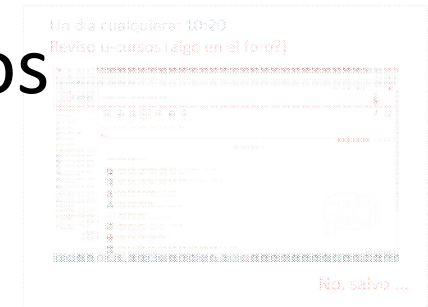
— *Bases de Datos Relacionales*

- El modelo más establecido

- El enfoque del curso

- Pero hablaremos brevemente de otros modelos

RUT	Nombre	Curso
...		
24.482.054-9	Aidan Hogan	CC3201
24.482.054-9	Aidan Hogan	CC5212
24.482.054-9	Aidan Hogan	CC6202



RUT	Nombre	Curso
...		
10.323.634-4	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC6202
12.323.792-8	Juan Ramírez	CC5212
...		

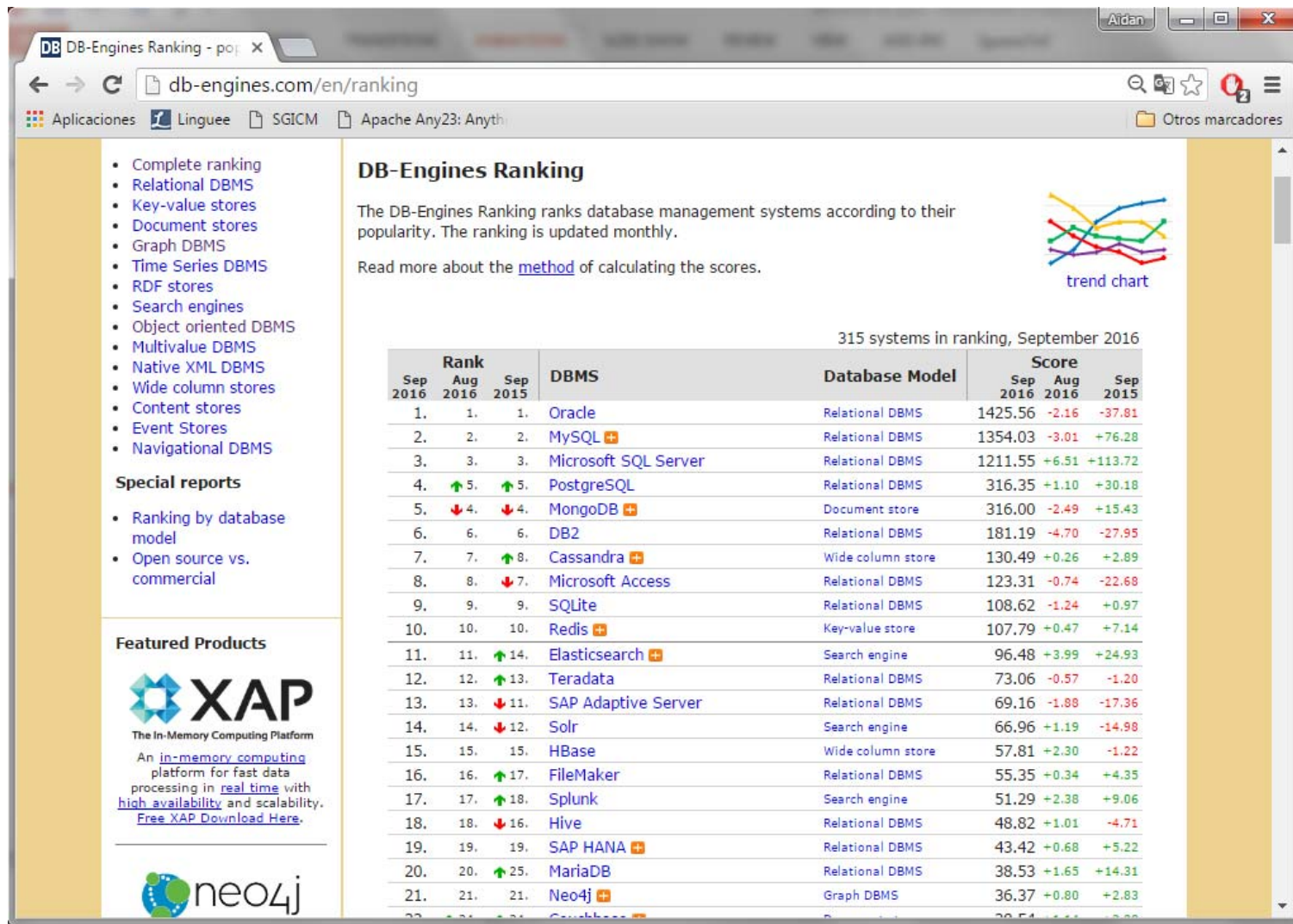
RUT	Nombre	Curso
...		
12.412.412-4	Sebastián Ferrada	CC3201
12.412.412-4	Sebastián Ferrada	CC5208
13.123.024-9	Daniel Hernández	CC6202
...		

RUT	Nombre
...	
CC3201	Bases de Datos
CC5212	Proc. Masivo de Datos
...	

RUT	Nombre	Eval	Nota
...			
10.323.634-4	CC3201	Lab1	6.0
10.323.634-4	CC3201	Lab2	4.5
12.323.792-8	CC6202	Examen	3.9
...			

**UNA DIVERSIDAD DE TIPOS DE
(SISTEMAS DE) BASES DE DATOS**

Los sistemas más utilizados en la práctica ...



<http://db-engines.com/en/ranking>

QUÉ VAMOS A APRENDER:

Una *introducción* a bases de datos:

- Hay tres tipos típicos de “usuarios” para un sistema de bases de datos:
 1. Usuarios finales
 2. Administradores del sistema
 3. Desarrolladores de un sistema
- Enfocaremos en el primer tipo
 - Incluye desarrolladores de *aplicaciones* de bases de datos
- Hablaremos menos de tipos dos y tres

En este curso, aprenderán

- Como se puede generalizar la consultación, la indexación, la gestión, etcétera, de datos
- Modelos de bases de datos
 - Con énfasis en el modelo relacional
 - Otros modelos: grafos, arboles, NoSQL ...
- Usar y manejar sistemas de bases de datos
 - Cargar datos, escribir consultas, actualizar datos

No aprenderán (específicamente)

- Rasgos específicos de todos los sistemas
- Como se puede implementar un sistema de basos de datos (en detalle)
- Minería de datos
- Sistemas distribuidos (en detalle)
- Datalog / lógica / “*deductive databases*”
- Temas de investigación

La estructura del curso

- Introducción / Motivación
- El Modelo Relacional
- La Algebra Relacional
- SQL (consultas)
- Diseño Relacional
- SQL (actualizaciones)
- Transacciones / Vistas
- Indexación / Optimización
- OLAP
- Otros Modelos: XML, Grafos, NoSQL

(Dado que es mi
primera vez
dictando este curso,
puede cambiar)

METODOLOGÍA DEL CURSO:

“Semi-flipped classroom”

- Los martes: una cátedra



“Semi-flipped classroom”

- Los jueves: una sesión interactiva:
 - Ejercicios escritos o laboratorios con notebooks
 - Trabajar en grupos para solucionar algo



Sesiones Auxiliares

- Los lunes: si perdemos una clase o si tratamos con temas difíciles, por ejemplo, tendremos una sesión auxiliar
 - No cada lunes!!
 - Pondré un mensaje en el foro al menos el jueves antes en anticipación

Material

- Subiré todas las diapositivas aquí después de cada cátedra:
 - <http://aidanhogan.com/teaching/cc3201-1-2016/>
- Las diapositivas servirán como el material canónico del curso
 - pero si quieren leer más, se recomienda:
 - “Database Management Systems”
 - Ramakrishnan / Gehrke, Third Edition

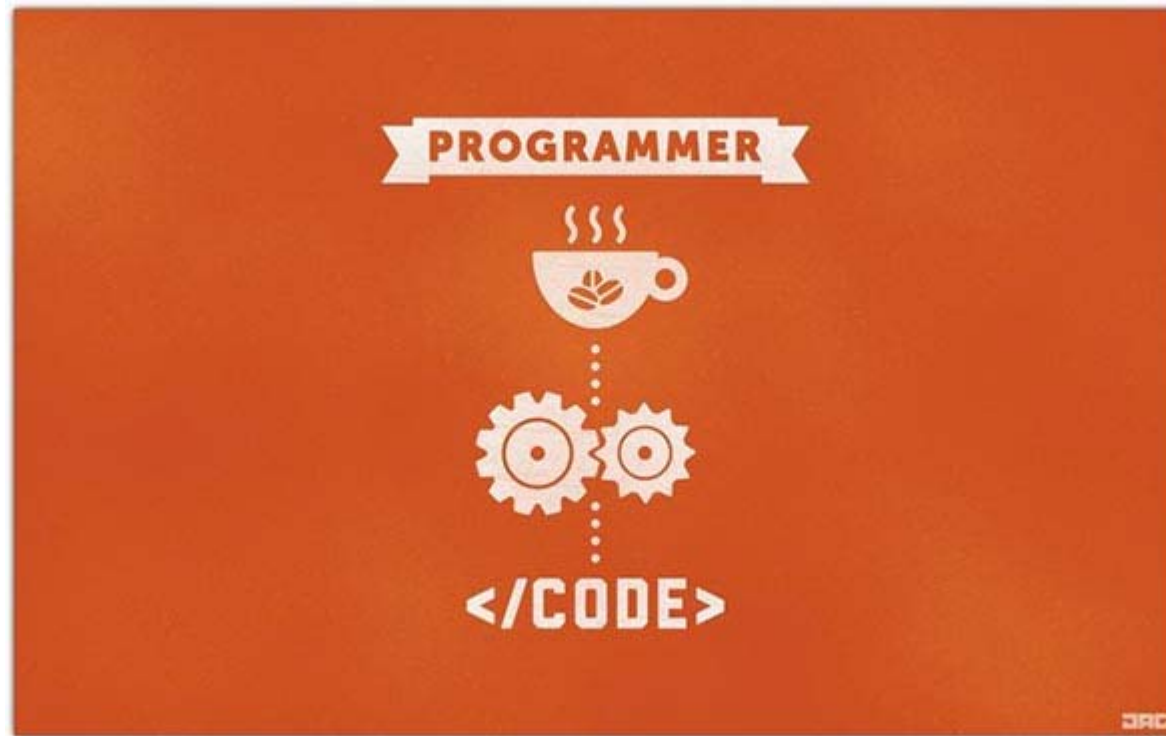
Evaluación del curso

- 50% tareas:
 - 40% ejercicios/laboratorios
 - 10% proyecto
- 50% exámenes:
 - 2 controles, 1 examen final

Este jueves ... laboratorio

- Vamos a implementar algo aquí
- Trabajaremos en grupos de dos o tres
- ¿Podrían llevar algunos notebooks por favor?

Sé programar en Java, sé programar en Python,
... podría programar algo sin problema!



... vamos a ver 😊.

Preguntas?

