MojoFacets

So, you have data and browser?

Dobrica Pavlinušić

`http://blog.rot13.org`

HULK, Knjižnica Filozofskog fakulteta u Zagrebu

Slobodni Festival 4, Čakovec, 2. srpnja 2010.

Goal

Create a **faceted data browser**, using web browser as a client with **data stored** on **server** in perl's **memory**.

Allow data **modification** while keeping **changes** to for recovery or replication.

Modify data and generate **reports** using simple **perl snippets** (it's in memory!)

Overview

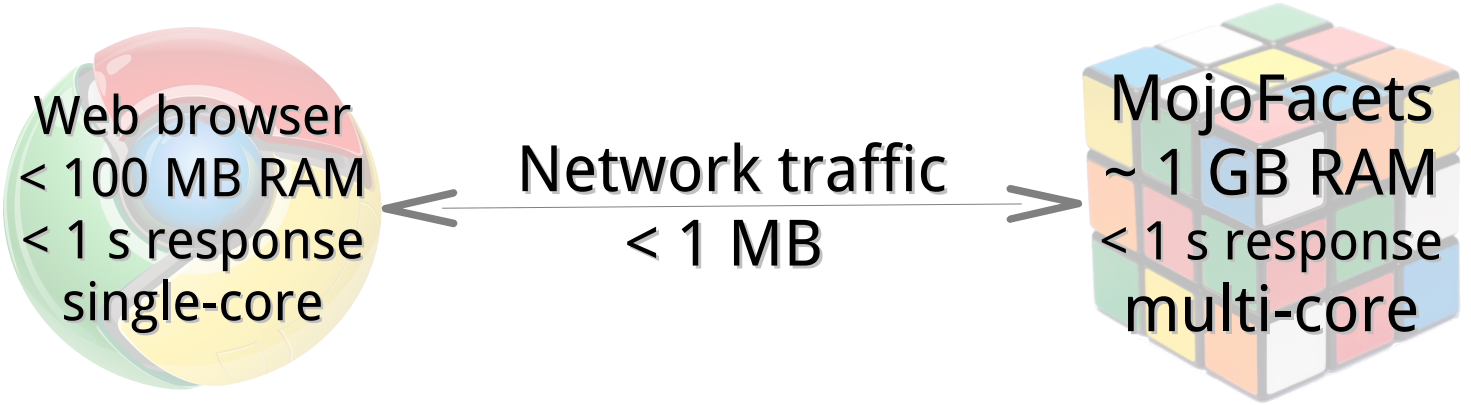
- Faceted browsing
- MojoFacets
  - Input formats
  - Faceted browsing, editing and plotting
  - Eval perl code over items or facets
- Google Chrome Development
  - Chrome Application
  - Speed Tracer

Facets?

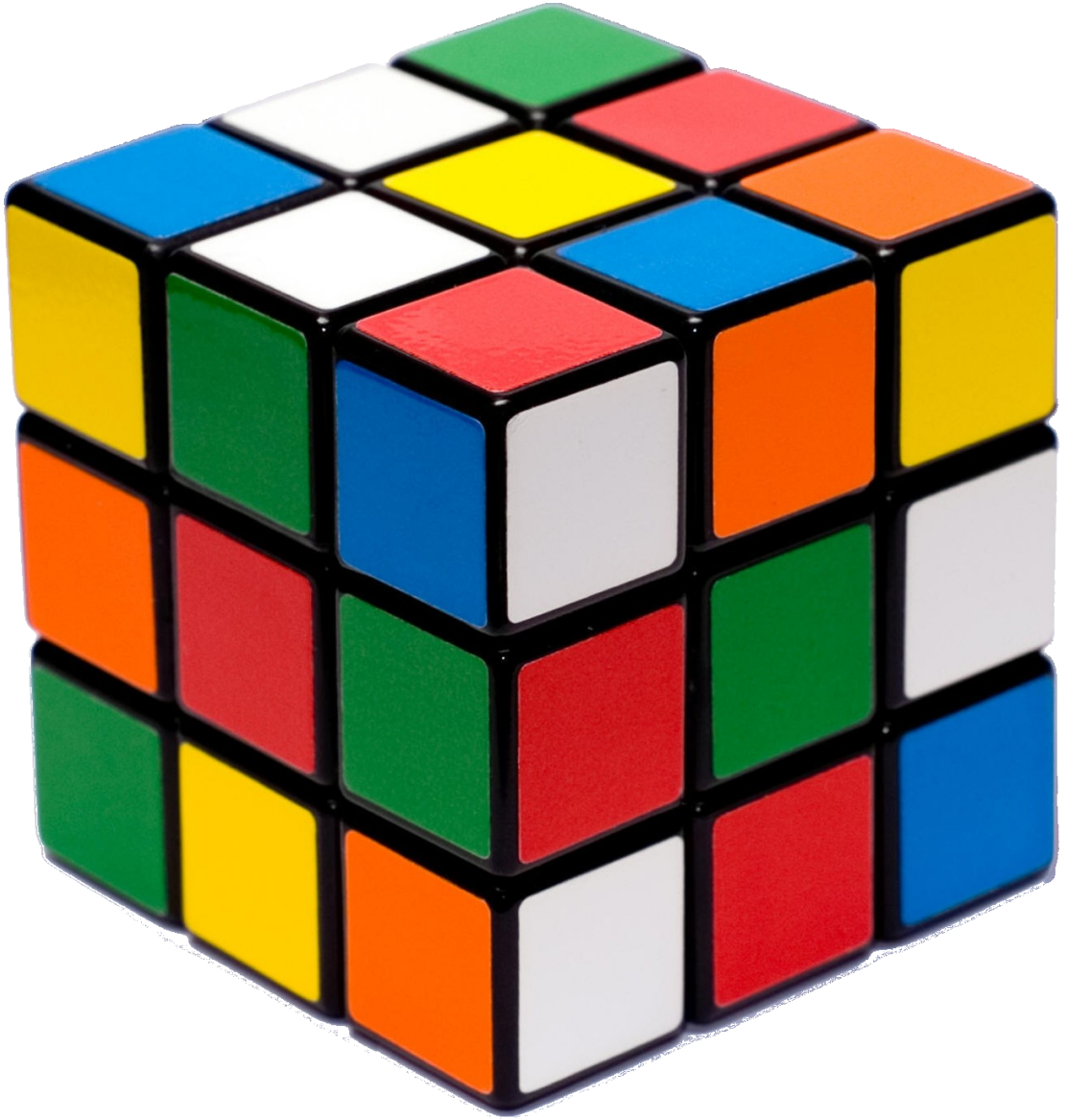
A **faceted classification** system allows the assignment of **multiple classifications** to an object, enabling the classifications to be ordered in multiple ways, rather than in a single, pre-determined, taxonomic order.

SIMILE Exhibit Widget – in-browser implementation for < 1000 items
http://www.simile-widgets.org/

Profile driven development



Assume modern browser (fast JavaScript, canvas), but transfer minimum HTML to reduce memory usage on device (Android) or CPU usage (netbook)



http://www.youtube.com/view_play_list?p=ED435109B7ABD426

MojoFacets

Web-based faceted browser with **server-side data in memory** for quick response

- Source data
  - JSON, CSV, html tables ...
- Server
  - Perl in-memory data with web server
  - Mojolicious, Devel::NYTProf
- Web application
  - jQuery UI (reorder columns)

Input formats

- JSON – Exhibit format with items
- TXT – pipe | delimited, multiline format
  - Plants for the Future
- CSV – cp1250 encoded
- Storable
  - Concateneted serialized perl objects
  - scripts/vmstat.pl – generate output

Server application

- Single user faceted browser
  - **No security, run on localhost!**
- Record all user **actions** and optionally replicate them to other **MASTER** instance
- Limit amount of HTML trasfered to browser
  - about:memory **< 100 Mb usage!**
  - **MAX_FACETS**
- Edit on double click with **changes** log
- Holds all data in memory as array of hashes
- Analyzes data lazily (numeric, unique)

Filter values = facet

- Subset of all values for a column in dataset
- Can be applied across different datasets with the same column name
- Selection of values using regex in browser
- Graph counts using canvas (< MAX_FACETS)
- Export to text file
- Copy/paste to spreadsheet (white ;)
- Eval perl code to filter or create agregates ($value, $count, $checked, $out)
- Multiple filters are combined and cached

Browsing and editing

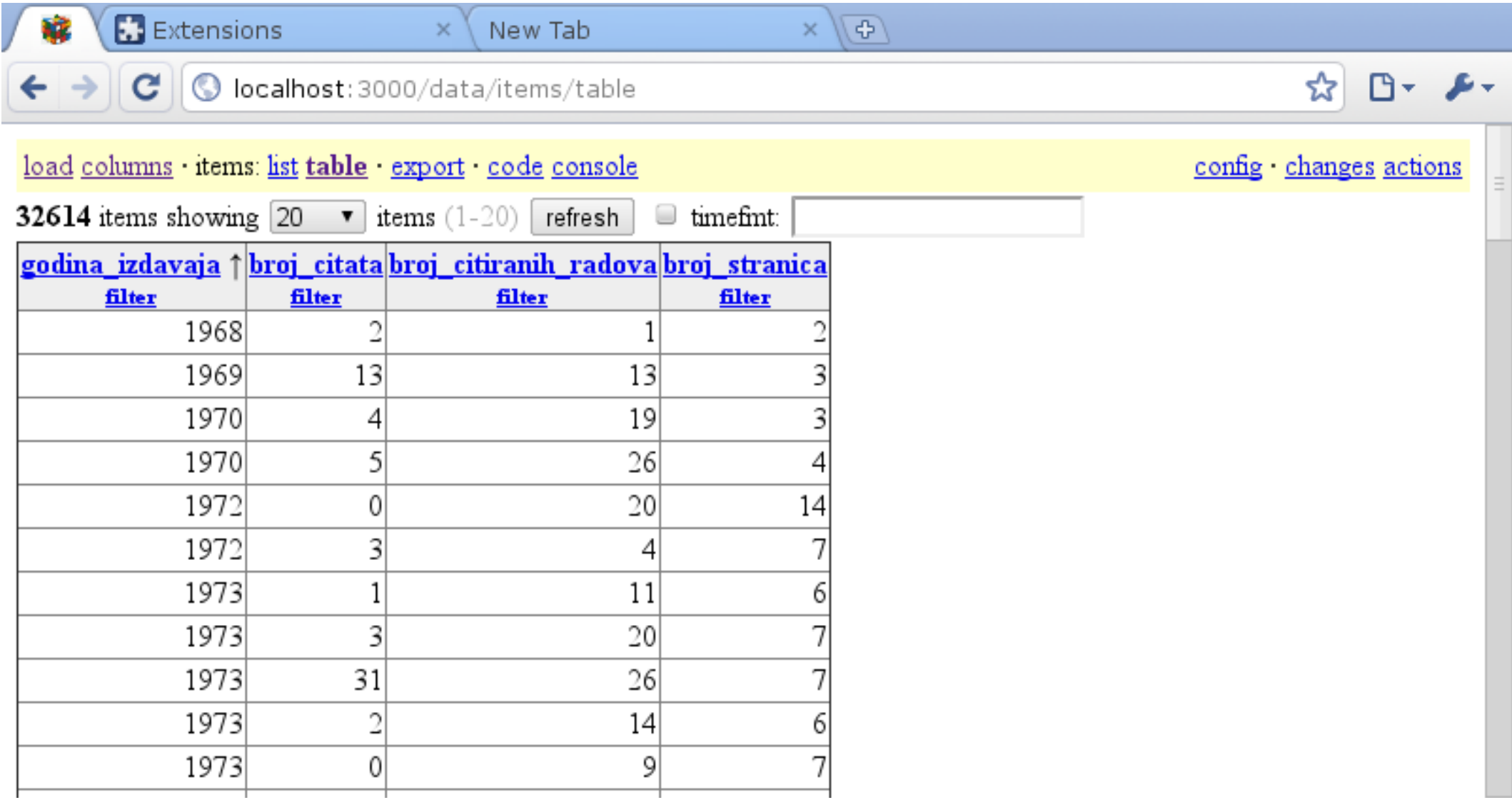
- Show items as table or list
- Graph all values using **gnuplot** server-side (and export data as side-effect)
- Edit any cell on double click
  - Repeatable values delimited by ¶
  - Create **changes** which can be applied to different datasets or new versions
- Load code snippets based on visible columns
  - `$row -> $update      # map`
  - `$out->{something}++ # collect`

Development
using
Google
Chrome

Google Chrome Application

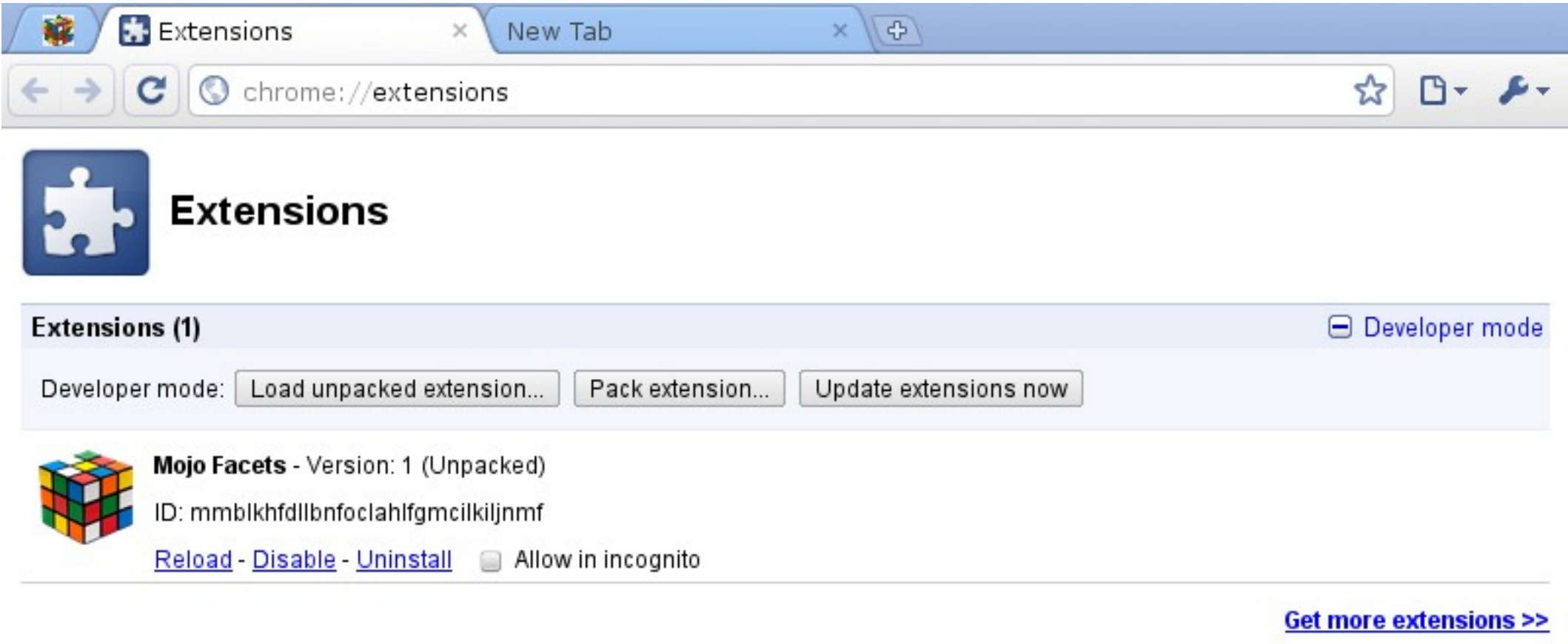
```
$ google-chrome --enable-apps

$ cat chrome-app/manifest.json
{
  "name": "Mojo Facets",
  "version": "1",
  "icons": { "24": "24.png", "128": "128.png" },
  "launch": {
    "web_url": "http://localhost:3000/"
  },
  "permissions": [],
  "web_content": {
    "enabled": true,
    "origin": "http://localhost"
  }
}
```

To get application tab...

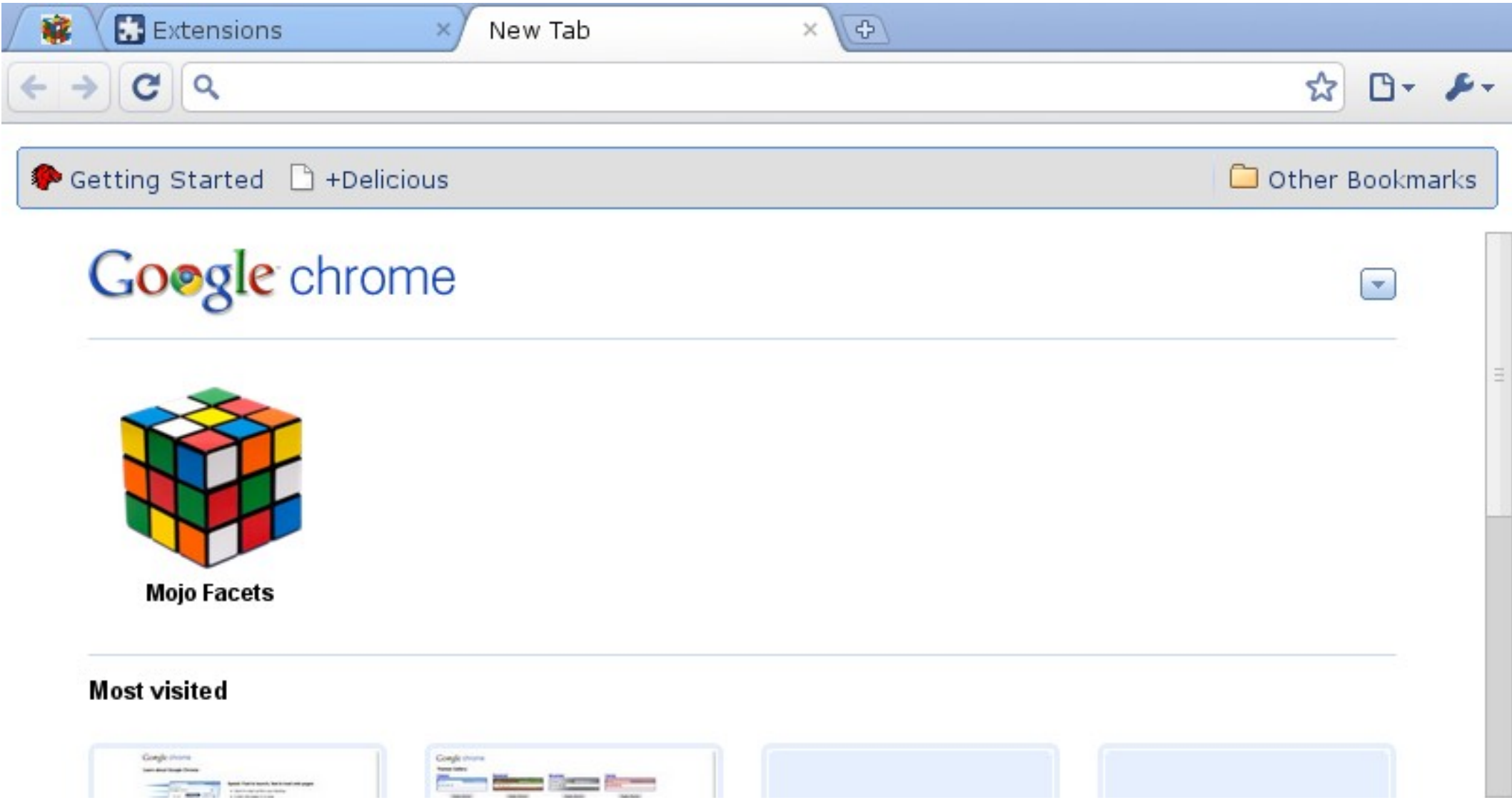


godina_izdavaja ↑ filter	broj_citata filter	broj_citiranih_radova filter	broj_stranica filter
1968	2	1	2
1969	13	13	3
1970	4	19	3
1970	5	26	4
1972	0	20	14
1972	3	4	7
1973	1	11	6
1973	3	20	7
1973	31	26	7
1973	2	14	6
1973	0	9	7

...load unpacked extension...

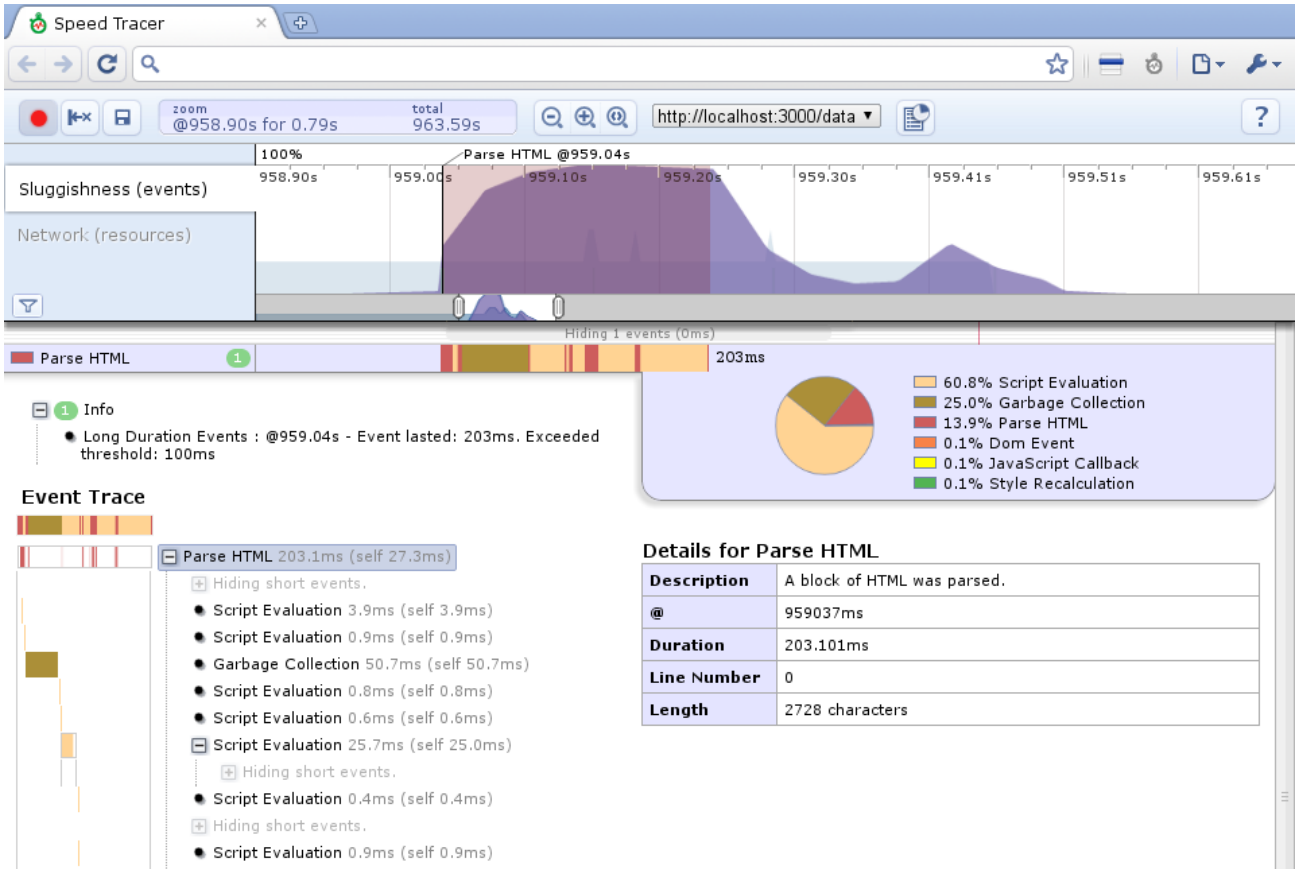


...and open new tab



Chrome Speed Tracer

```
$ google-chrome --enable-extension-timeline-api
http://code.google.com/p/speedtracer/
```



Find out more!

Blog

http://blog.rot13.org/mojofacets

github

http://github.com/dpavlin/MojoFacets

Project updates

https://www.ohloh.net/p/MojoFacets

http://blog.rot13.org