



An API for Your Data

David Brennan, PhUSE 2015



Background

- Tables, tables, tables...
- Open-source tools
 - Web development frameworks
 - Javascript libraries
 - html/css frameworks
- Data in the browser
- Want to build simple tools / interactive presentations / visualisations



Background cont'd...

- In general we use denormalised data tables
 - e.g. SDTM + ADAM...
 - ...could use a relational database...
 - ...just don't need joins, only need to query data!
- The Plan:
 - Import data into relational database
 - Provide simple query API to data
 - Do stuff with data in the browser...



API??

- **API** = **A**pplication **P**rogramming **I**nterface
- Specifically, a web API...
- Wikipedia –
 - *When used in the context of **web** development, an **API** is typically defined as a set of Hypertext Transfer Protocol (**HTTP**) **request messages**, along with a definition of the structure of **response messages**, which is usually in an Extensible Markup Language (XML) or JavaScript Object Notation (**JSON**) format*
- Or...
 - **A bunch of URLs**

Patient Profile-Style Display

Simple Patient Profile - Chromium

Simple Patient Profile x

davbre.github.io/mira-examples/ex6.1_simple_profile.html#info-box-toggle

Index

-- select a subject ID --

- ABC001-101-01
- ABC001-101-02
- ABC001-101-03
- ABC001-101-04
- ABC001-101-05

Hide Menu

Show/hide info box

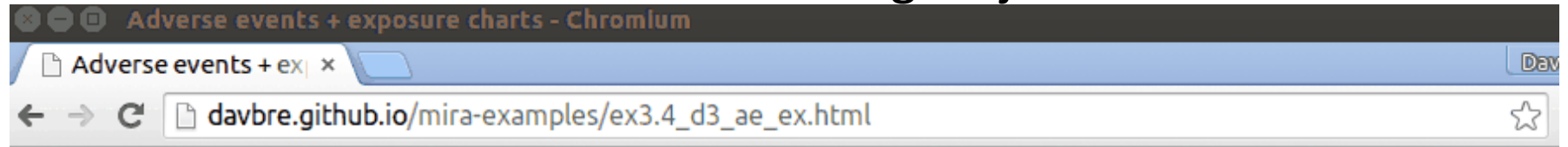
Patient Summary

Exposure

Adverse Events

Laboratory Results

AE + EX Chart using d3.js



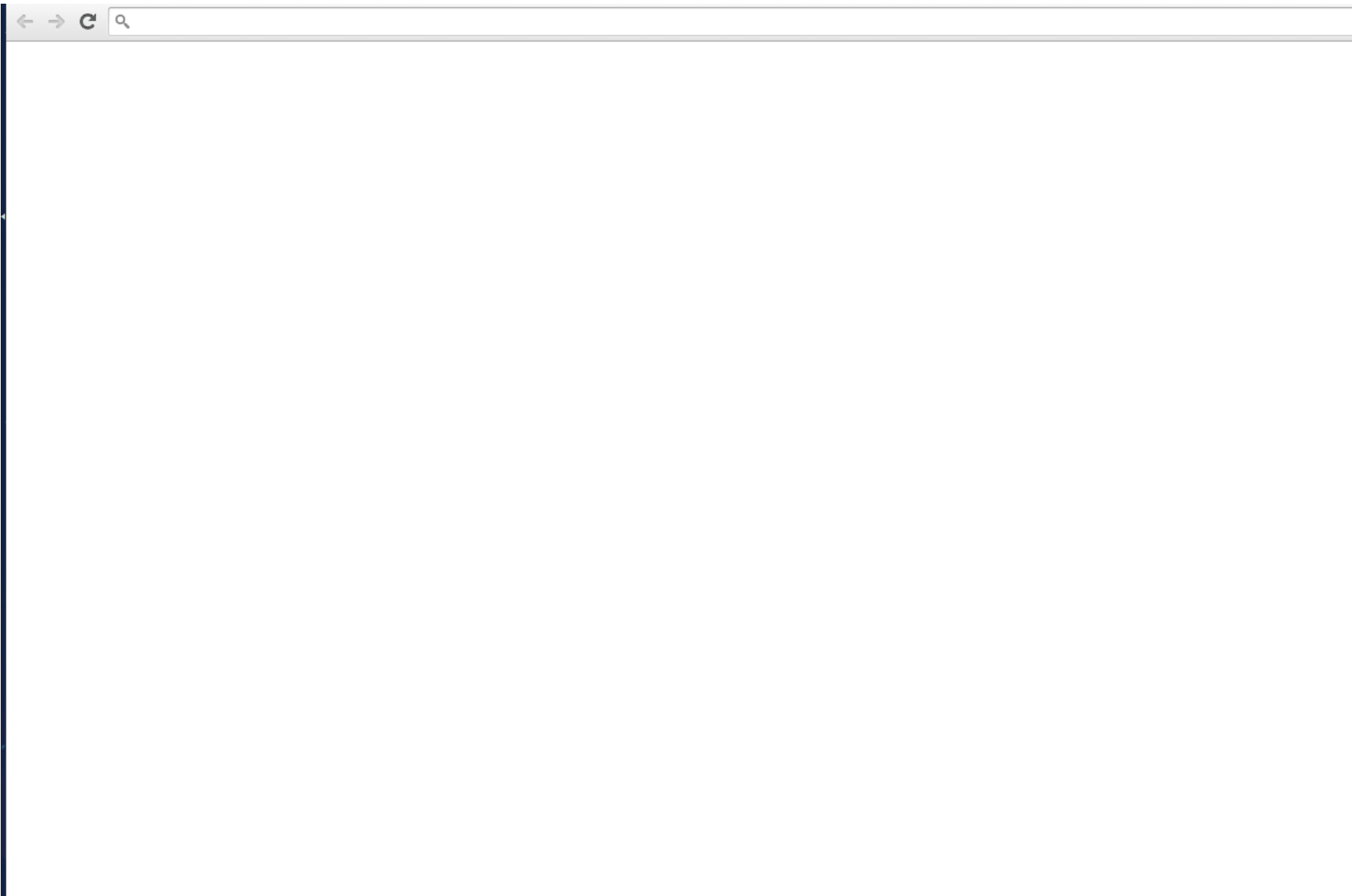
[Index](#)

-- select a subject ID --

- ABC001-101-01
- ABC001-101-02
- ABC001-101-03
- ABC001-101-04
- ABC001-101-05



AE Dashboard using crossfilter.js + dc.js





Web API Examples

- Ebay, Amazon, Twitter, Facebook, Google maps / analytics, LinkedIn...
- Ireland Revenue Customs and Excise API
- ONS UK – OpenAPI
- United States Census Bureau – various APIs
- Berkeley Ecoinformatics Engine – natural history data
- **openFDA : Open-source APIs for FDA drug, device and food data**



Why an API??

- Gives simple building blocks for data presentations / visualisations
- Want data in the browser
 - Lots of great libraries
 - jQuery, d3, c3, google-charts, crossfilter, dc, DataTables...
 - + many more
 - Easy to get started → browser + text-editor
 - css / css-frameworks for presentation
- Separation of data from its presentation



Introducing Mira

- Web application

<https://github.com/davbre/mira>

← has link to demo

<http://davbre.github.io>

← has examples of
interacting with an API

- Upload tables of data + metadata
- You get a HTTP API to uploaded tables
- Do interesting things in browser using API, e.g...



- Dashboards / Mash-ups
- Interactive / dynamic charts
- Presentations such as patient profiles, TFLs and linking them up
- Data-viewing – multiple tables at once
- Saving state - bookmark data views
- ...???



Projects | Mira - Chromium

Projects | Mira

David

← → ↻ 46.101.208.152

☆ ≡

MIRA

Projects

Logged in as david.n.brennan@gmail.com. | Logout

New project

Project ID	Project Name	Description			
5	Dummy SDTM Data	Some dummy clinical trial data in SDTM-like format + some AE reports	API details	Edit Project	delete
4	World GDP	World GDP figures - http://data.okfn.org/data/core/gdp.datapackage.json - changed year type from "date" to "number" as it only gives the year.	API details	Edit Project	delete
3	Populations	Populations - http://data.okfn.org/data/core/population. Changed "Year" type from "date" to "integer" in order to upload to postgresql without errors.	API details	Edit Project	delete
2	Europe, list of refugee deaths 1993-2012	List of deaths of refugees in europe and on europes borders. From http://data.okfn.org/data/mihi-tr/list-of-deaths-1993-2012.	API details	Edit Project	delete
1	Airport Codes	Airport codes from http://ourairports.com/data/	API details	Edit Project	delete



Mira - Chromium

Mira x David

46.101.208.152/projects/5/api-details

MIRA

API Endpoints

Description	HTTP Method	API Endpoint
All projects	GET	http://46.101.208.152/api/projects
This project	GET	http://46.101.208.152/api/projects/5
Project tables	GET	http://46.101.208.152/api/projects/5/tables
Project datapackages	GET	http://46.101.208.152/api/projects/5/datapackages
Tables		
adae	GET	Detail http://46.101.208.152/api/projects/5/tables/adae Columns http://46.101.208.152/api/projects/5/tables/adae/columns Data http://46.101.208.152/api/projects/5/tables/adae/data
adlb	GET	Detail http://46.101.208.152/api/projects/5/tables/adlb Columns http://46.101.208.152/api/projects/5/tables/adlb/columns Data http://46.101.208.152/api/projects/5/tables/adlb/data
adlb_pc	GET	Detail http://46.101.208.152/api/projects/5/tables/adlb_pc Columns http://46.101.208.152/api/projects/5/tables/adlb_pc/columns Data http://46.101.208.152/api/projects/5/tables/adlb_pc/data
ae	GET	Detail http://46.101.208.152/api/projects/5/tables/ae Columns http://46.101.208.152/api/projects/5/tables/ae/columns Data http://46.101.208.152/api/projects/5/tables/ae/data

index

- Summary of Serious Adverse Events
- Summary of Serious Adverse Events - Female Subjects
- Summary of Serious Adverse Events - Male Subjects
- Summary of Severe Adverse Events
- Summary of Severe Adverse Events - Female Subjects
- Summary of Severe Adverse Events - Male Subjects

[illegible]

Application using Mira API to browse data + bookmark views

The screenshot shows a web browser window titled "Mira Bookmarks - Chromium" with a single tab labeled "Mira Bookmarks". The address bar displays "localhost:3001/api_urls/1/bm-viewer". The application interface features a dark blue sidebar on the left. The main content area has a top navigation bar with a "Home" link and a user status "Logged in as admin. Edit profile | Logout". Below this, there are three sections: "Projects", "Tables", and "New Bookmark". The "Projects" section contains a scrollable list with items "Populations", "Census 2010", and "Dummy SDTM". The "Tables" section is currently empty. At the bottom of the main content area, there are two buttons: "< Hide Bookmarks" and "✕ Clear All Filters".

Mira Bookmarks - Chromium

Mira Bookmarks x

localhost:3001/api_urls/1/bm-viewer

Home

Logged in as **admin**. [Edit profile](#) | [Logout](#)

Projects

Populations
Census 2010
Dummy SDTM

Tables

New Bookmark

< Hide Bookmarks ✕ Clear All Filters



Tools + Technology

- Ruby on Rails (web development framework)
 - PostgreSQL (relational database)
 - Upload CSV files
+
metadata
- ➔ HTTP API



Why CSV?

- IBM Fortran (level G) compiler under OS/360 supported CSV files in 1967!
- Simple, common “standard” → easy to work with
- Supported by lots of tools
- **Tablular Data Packages** = CSV files + their metadata in **datapackage.json**



Tabular Data Packages

- A simple data publishing format promoted by the Open Knowledge Foundation:

<http://data.okfn.org/doc/tabular-data-package>

- Data is stored in CSV
- Metadata in a file called **datapackage.json**



datapackage.json

```
{
  "name": "Study ABC001",
  "resources": [
    {
      "path": "ADAE.csv",
      "dialect": { "delimiter": "," },
      "schema": {
        "fields": [
          { "name": "USUBJID", "type": "string" },
          { "name": "AETERM", "type": "string" },
          { "name": "AESTDT", "type": "date" },
          { "name": "AESTDY", "type": "integer" },
          { ...etc etc } ]
        }
      },
    {
      "path": "ADCM.csv",
      ...etc etc
    }
  ]
}
```



HTTP Requests

Mira uses the term “projects” to describe groups of uploaded files. Assuming an ID number of 123 has been assigned:

GET info on uploaded tables:

<http://.../api/projects/123/tables>

GET info on “adae” columns:

<http://.../api/projects/123/tables/adae/columns>

GET a page of data from “adae”:

<http://.../api/projects/123/tables/adae/data>

GET rows from “adae” table where subject = ABC001-101-03:

http://.../api/projects/123/tables/adae/data?usubjid_eq=ABC001-101-03



Another HTTP Request

- Field-value pairs...

GET serious adverse events, not mild, “aedecod” contains “Headache”, “aebodsys” begins with “Nuero” and “aeendt” is missing. Order by “siteid”, “usubjid” and descending “aestdt”. Return 30 rows per 'page' and the 3rd page:

```
http://.../api/projects/123/tables/adae/data?  
  aeser_eq=Y  
  &aesev_ne=Mild  
  &aedecod_contains=Headache  
  &aebodsys_begins=Neuro  
  &aeendt_blank  
  &order=siteid,usubjid,aestdt:desc  
  &per_page=30  
  &page=3
```



JSON Response

Request: GET <http://46.101.208.152/api/projects/5/tables/sv/data>

Response:

```
{data:
  [
    {id: 1,
      studyid: "ABC001",
      usubjid: "ABC001-101-01",
      visitnum: 1,
      visit: "SCREENING",
      visitdy: -15,
      svstdtc: "2014-05-28",
      svendtc: "2014-05-28",
      svstdy: -15,
      svenidy: -15
    }, {
      id: 2,
      studyid: "ABC001",
      usubjid: "ABC001-101-01",
      visitnum: 2,
      ...etc etc
```



Mira...take a look...

Repository: <https://github.com/davbre/mira>

Mira is open-source

Links to live instance of Mira providing an API to some dummy SDTM data

Some information on setting up your own local instance

Examples: <http://davbre.github.io>

Some examples of interacting with a Mira API

Contact: gmail ➔ david.n.brennan