INTERACTION DESIGN

DATA TYPES

Data Visualisation HS21

DATA TYPES

Structured	Unstructured
Pre-defined data model / annotated	Not organized in a pre-defined manner
Suited for a certain processing task	Not suited for the processing task at hand
Often tables / databases	Often text-heavy / images / music

DATA TYPES

Propietary*	Textual**
MP3	PDF
Spotify Tracks	JSON
PSD	XML
AEP	CSV

*Restricted, secret
**Open, public

DATA TYPES

XML

(Extensible Markup Language)

Describes how data should be organized. Uses tags that are readable, which makes manually traversing the data really easy.

```
<entry>
    <file_name>DataVis</file_name>
    <date>18-11-2021</date>
    <author>Paulina</author>
    <file_size>17kb</file_size>
</entry>
```

JSON

(JavaScript Object Notation)

Compact but *does* include formatting characters
Allows for more complex data structures, including objects and arrays

```
{

“file name”: "datavis.txt",
"date" : "18-11-2021",
"author": “paulina",
"file size": “17kb",

}
```

CSV

(Comma Separated Value)

Uses coma to separate values
Minimizes the extra formatting characters

```
  File name, date, author, file size
datavis.txt, 18-11-2021, Paulina, 17kb
```

JSON vs CSV

JSON	CSV
JSON is used for storing and exchanging the data.	CSV is delimiting text that uses the comma to separate the file.
Each object can have different fields, and the field order is not significant	All the records should have the same fields, and be in the same order.
It is less compact as compared to the CSV file	The CSV files are more compact than other file formats
JSON has a lot of scalability in terms of adding and editing the content	CSV lacks scalability
Can contain nested structures	Is based on two dimensional and tabular format
Uses JavaScript data types	Requires external metadata to identify data types.
Great for big amount of data, especially while using API	Great for handling small amount of data

INTERACTION DESIGN

PREPARING DATA

Data Visualisation HS21

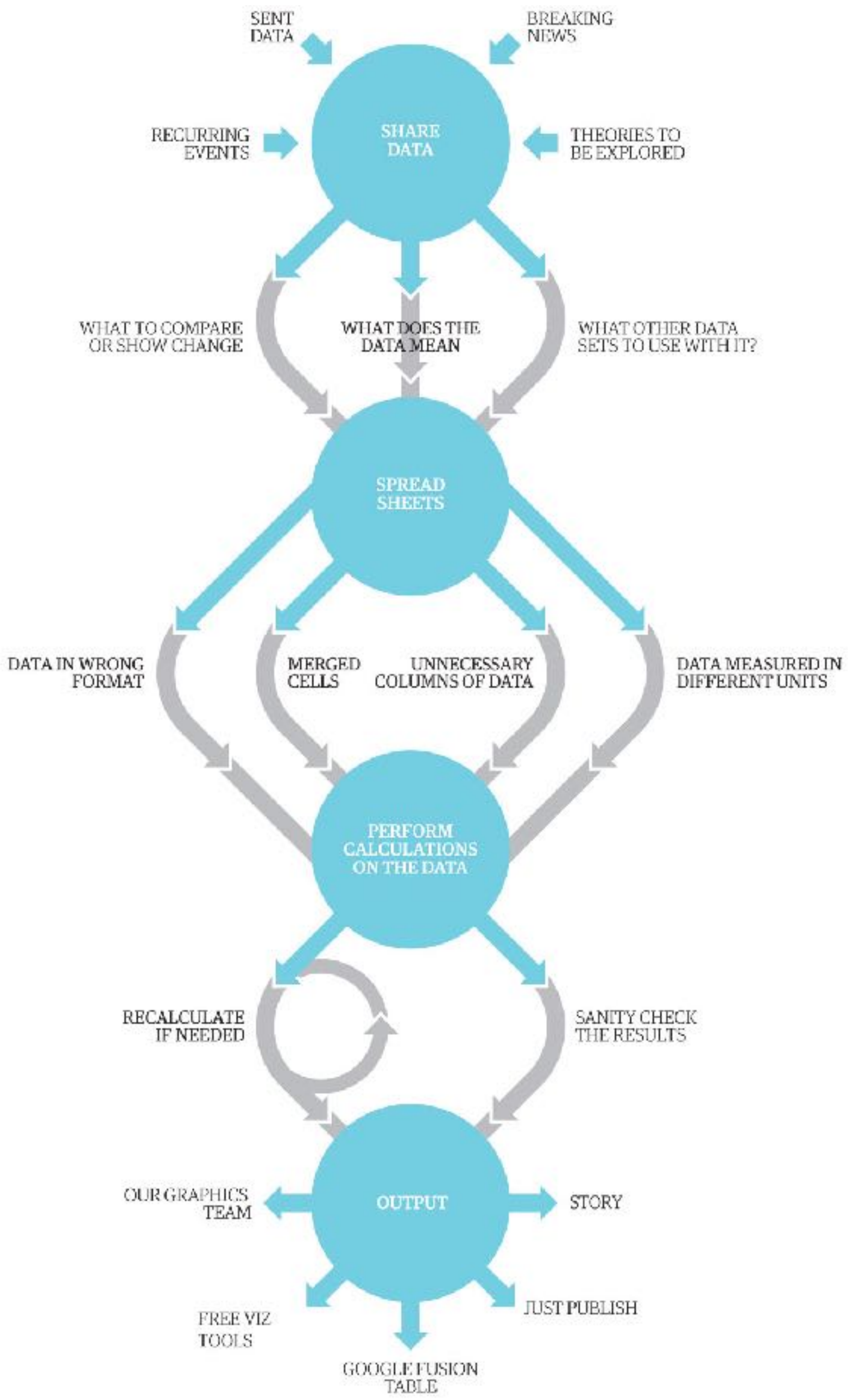
Data search

Often, data are publicly available, but not structured enough. E.g. in a table on a website, but not as a downloadable file format.

- We then need to scrape the website for the data.
- Often, this is done in an automated fashion, either with self-programmed bots or software tailored for this task.
- Even more often, it has to be made manually

https://www.niemanlab.org/2014/12/a-wave-of-p-r-data/

DATA WORKFLOW



SENT DATA

BREAKING NEWS

RECURRING EVENTS

SHARE DATA

THEORIES TO BE EXPLORED

WHAT TO COMPARE OR SHOW CHANGE

WHAT DOES THE DATA MEAN

WHAT OTHER DATA SETS TO USE WITH IT?

SPREAD SHEETS

DATA IN WRONG FORMAT

MERGED CELLS

UNNECESSARY COLUMNS OF DATA

DATA MEASURED IN DIFFERENT UNITS

PERFORM CALCULATIONS ON THE DATA

RECALCULATE IF NEEDED

SANITY CHECK THE RESULTS

OUR GRAPHICS TEAM

OUTPUT

STORY

FREE VIZ TOOLS

JUST PUBLISH

GOOGLE FUSION TABLE

https://simonrogers.net/2013/01/27/a-data-journalism-workflow/

Data search

https://data.stadt-zuerich.ch/

https://ourworldindata.org/

https://opendata.swiss/en

https://www.crowtherlab.com/maps/#/

https://ec.europa.eu/eurostat/de/web/main/data/database

https://hack.opendata.ch/

https://www.research-collection.ethz.ch/handle/20.500.11850/31

https://hub.arcgis.com/search

https://gisgeography.com/openstreetmap-download-osm-data/

https://disc.gsfc.nasa.gov/

Exercise 1

In groups of 2 try to find data that can answer the task you received in the Data Visualisation module:

- Present and live demo us where you found the data
- Write down the website, the data format and the tools you used

https://click-that-hood.com/?city=zurich-city

Data quality

- Values are missing
- Date formats are inconsistent
- Spelling is inconsistent
- Text is garbled (“encoding hell”)
- Data type is inconsistent

Data quality

OpenRefine

A free, open source, powerful tool for working with messy data

JSON Formatter

Offered by: callumlocke.co.uk

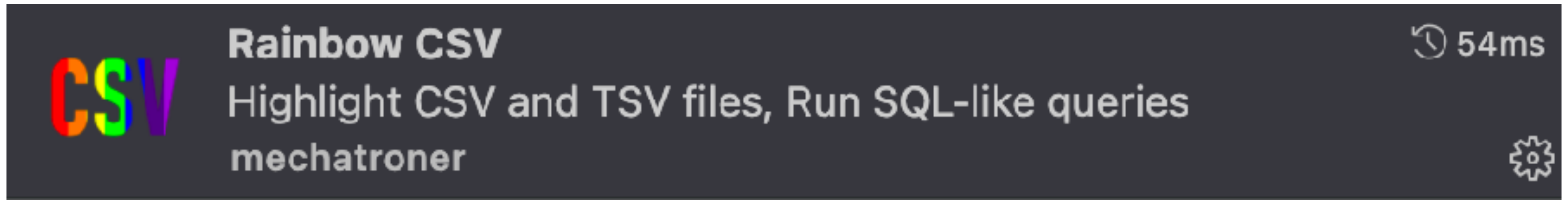
★★★★★ 1,761 | Developer Tools | 1,000,000+ users

{JSON formatter}

CSV**JSON**

Sponsored by **Flatfile.io**

Useful Tools (VSCode)

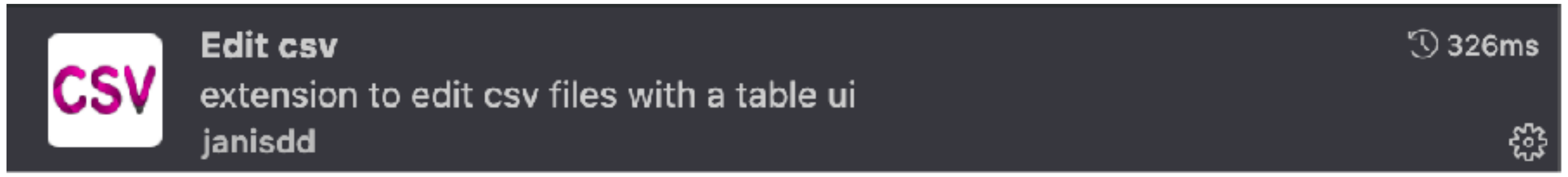


Rainbow CSV

Highlight CSV and TSV files, Run SQL-like queries

mechatroner

54ms

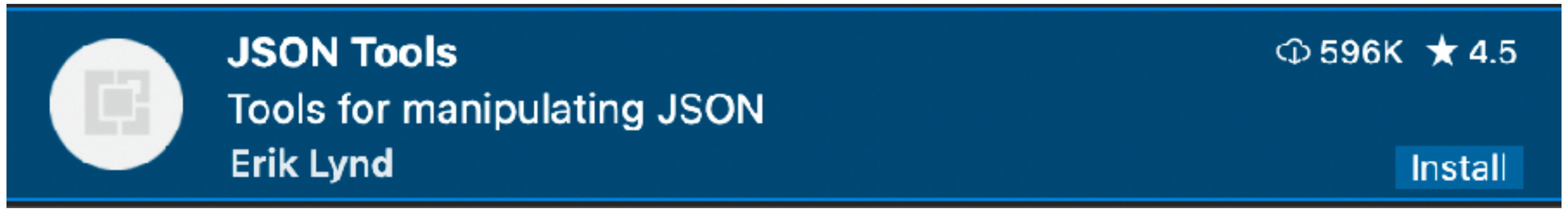


Edit csv

extension to edit csv files with a table ui

janisdd

326ms

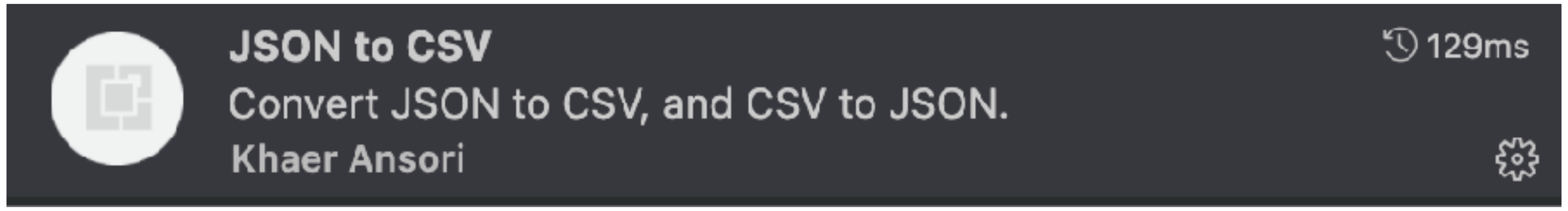


JSON Tools

Tools for manipulating JSON

Erik Lynd

596K ★ 4.5



JSON to CSV

Convert JSON to CSV, and CSV to JSON.

Khaer Ansori

129ms

Data conversion

future_cities_data.csv ↔ future_cities_data.json

INTERACTION DESIGN

HANDLING DATA IN P5.JS

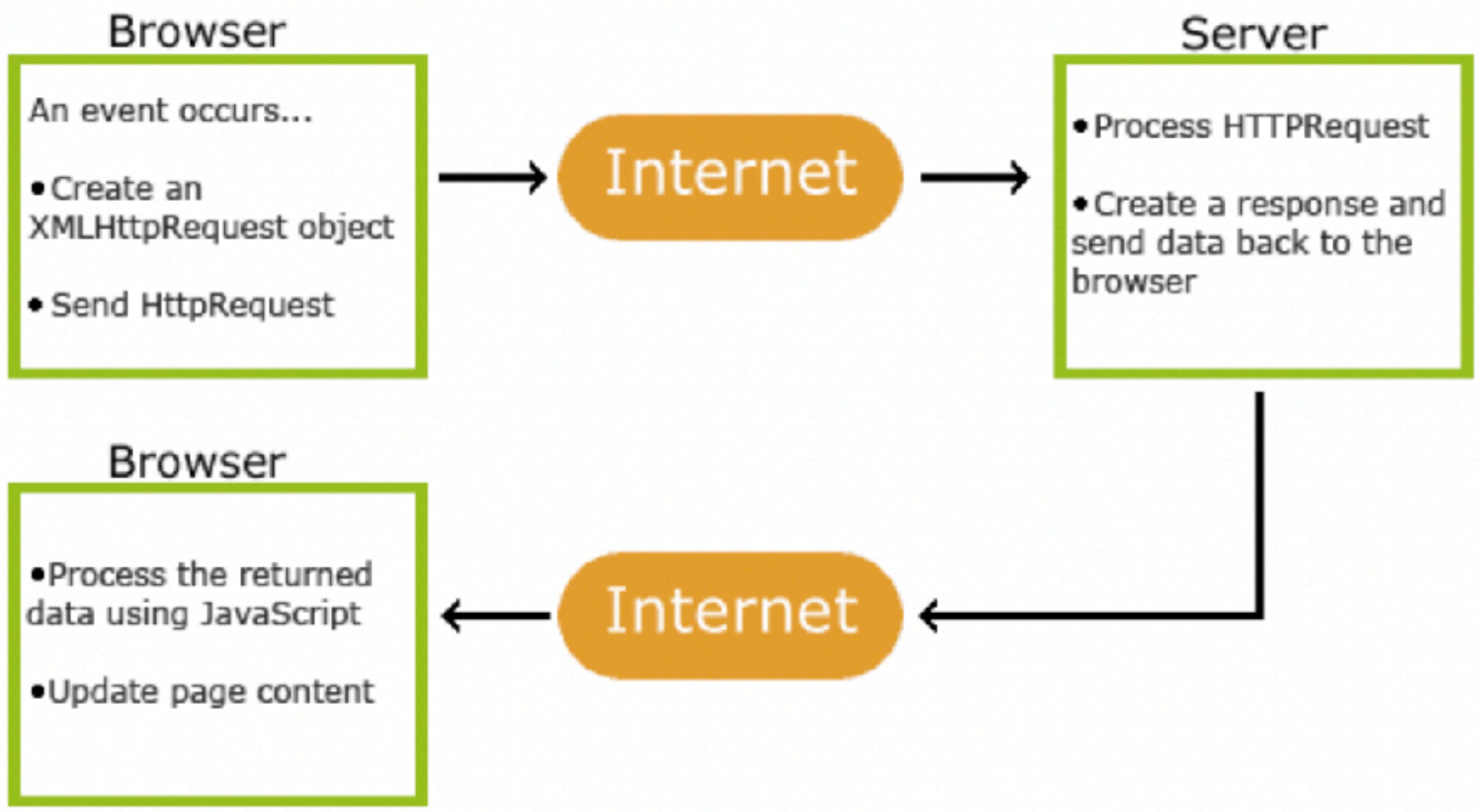
Data Visualisation HS21

AJAX

JS is "single threaded and synchronous", meaning everything runs in order that it's written in the file. **However**, JS also makes use of "asynchronous" functions to make the program flow faster.

AJAX

AJAX (Asynchronous JavaScript and XML) is a technique for communicating with a server and dynamically altering a page without leaving the page. It is made possible with the XMLHttpRequest object, a built-in feature of a web-browser.



Browser

An event occurs...

- Create an XMLHttpRequest object
- Send HttpRequest

Internet

Server

- Process HTTPRequest
- Create a response and send data back to the browser

Internet

Browser

- Process the returned data using JavaScript
- Update page content

AJAX

p5.js has a variety of methods for using AJAX to retrieve files from a server. These include:

- **loadStrings** - loads .txt files
- **loadJSON** - loads .json files
- **loadXML** - loads .xml files
- **loadTable** - loads .csv files

p5.Table

```
var myTable = loadTableloadTable(filename, [header], [callback], [errorCallback])
```

- `.getRowCount()`: evaluates to the number of rows in the file
- `.getNum(i, "foo")`: evaluates to the value in the cell in row `i` with column name `foo`
- `.getColumn("foo")`: evaluates to an array of all the values in column `foo`
- …and many others (reference)

p5.Table

```
var myTable = loadJSON(path, callback, [errorCallback])
```

- `.getRowCount()`: evaluates to the number of rows in the file
- `.getNum(i, "foo")`: evaluates to the value in the cell in row `i` with column name `foo`
- `.getColumn("foo")`: evaluates to an array of all the values in column `foo`
- …and many others (reference)

Exercise 2

In groups of 2 use the default p5 **loadTable** or **loadJSON** function to load the data from **exercise 1** and visualise a correlation between two input fields.

AJAX

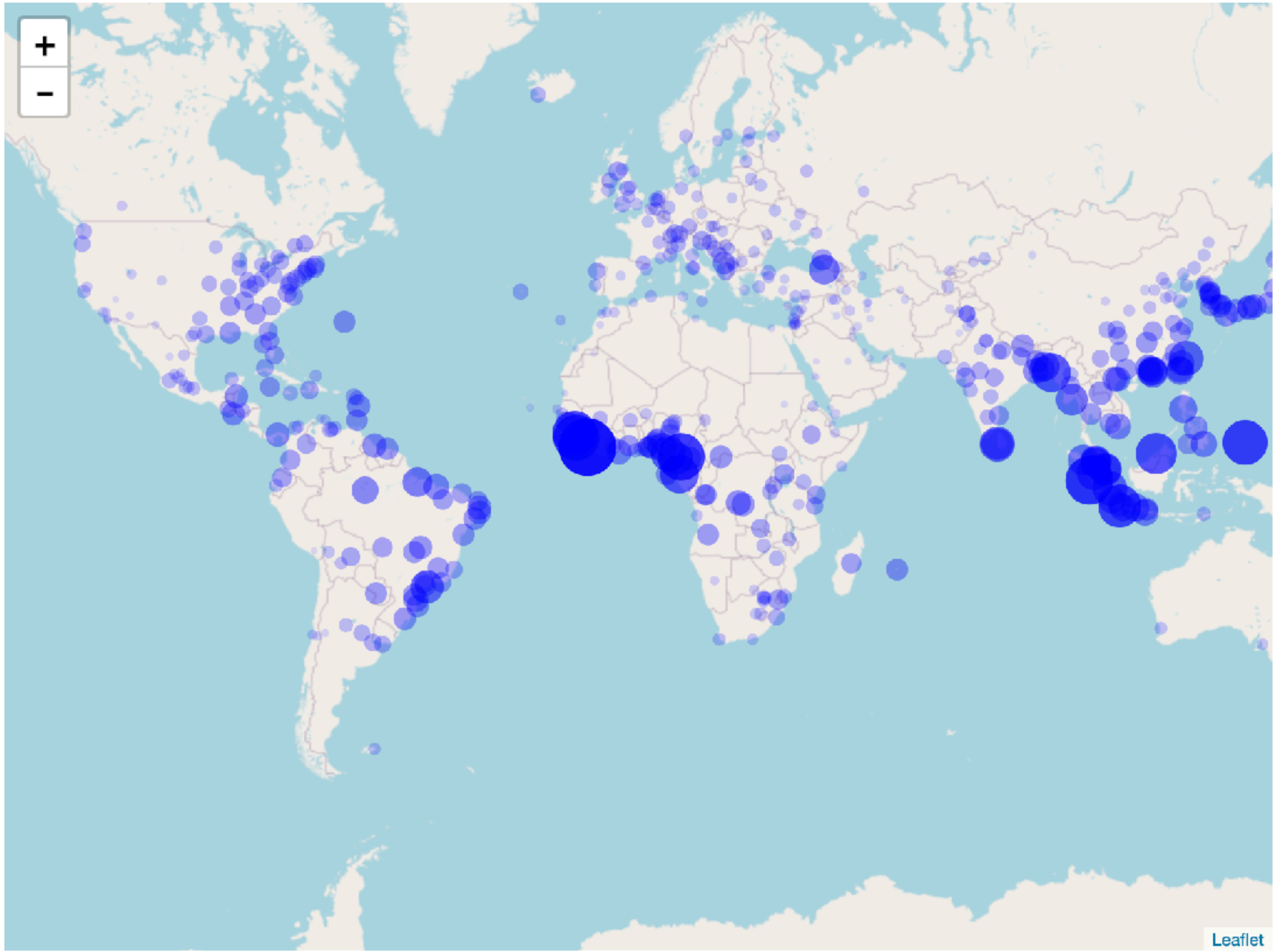
A **callback** is a function that is passed to another function as a parameter, and called by that other function. A **callback function** is useful when working with asynchronous functions because it allows us to specify some code to execute after the first asynchronous task has completed.

INTERACTION DESIGN

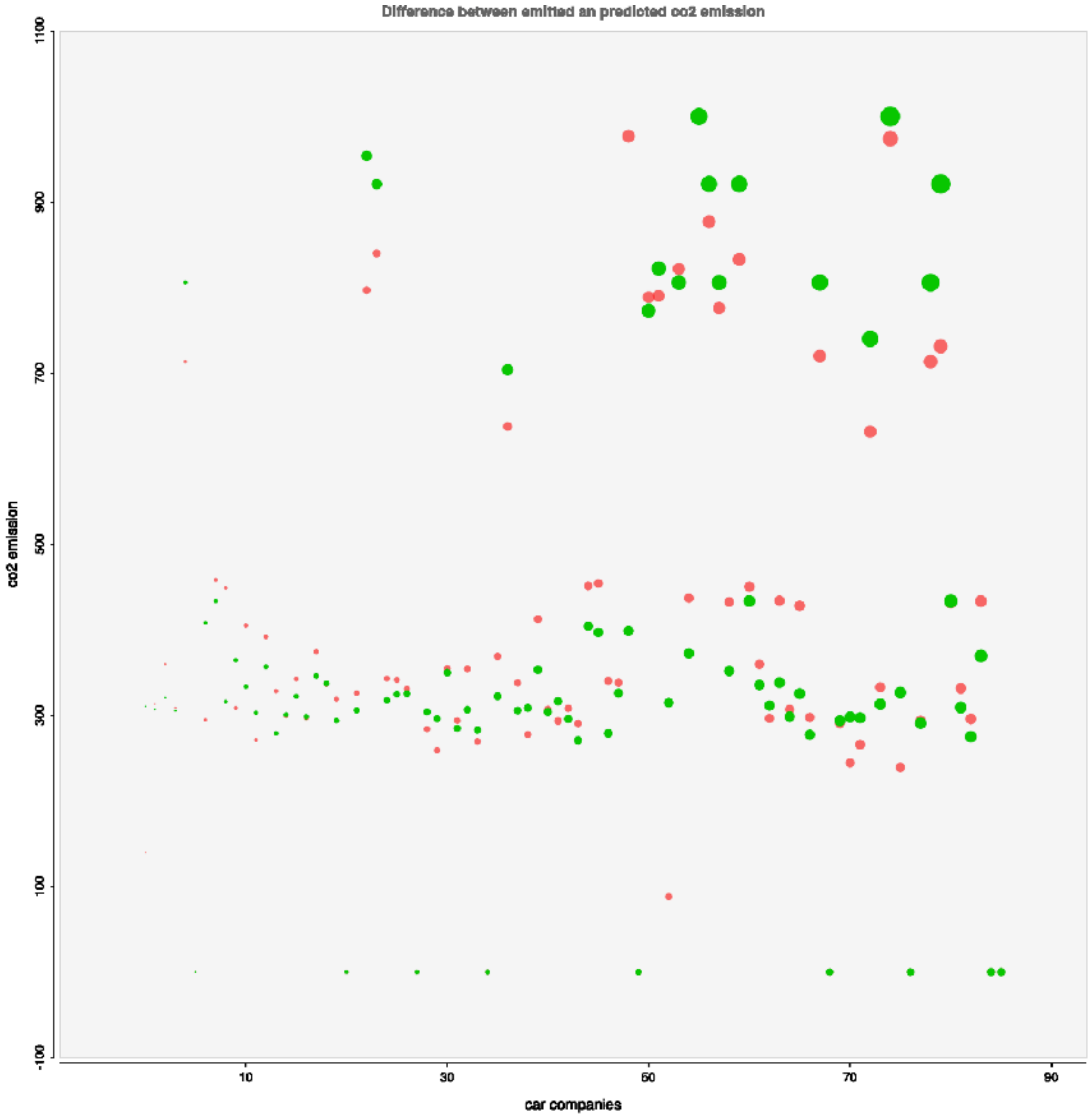
DATVIZ LIBRARIES

Data Visualisation HS21

Mappa.js



grafica.js


Difference between emitted an predicted co2 emission
1100
900
700
500
300
100
-100
co2 emission
10
30
50
70
90
car companies

D3.js

d3.js is a JavaScript library for manipulating documents based on data, where the data is mapped to HTML elements. So d3.js is heavily tied to HTML document structure (or DOM) for accessing those elements and creating, removing or modifying them based on data.

HTTPS://BOST.OCKS.ORG/MIKE/D3/WORKSHOP/#0

HTTPS://SCIUTOALEX.GITHUB.IO/P5-D3-COOKBOOK/

HTTPS://GITHUB.COM/SEPANS/P5D3

HTTPS://GITHUB.COM/D3/D3/BLOB/MAIN/API.MD#GEOGRAPHIES-D3-GEO

INTERACTION DESIGN

GEODATA

Data Visualisation HS21

.geojson

GeoJSON is a way of representing geographic data and consists of the following different parts:

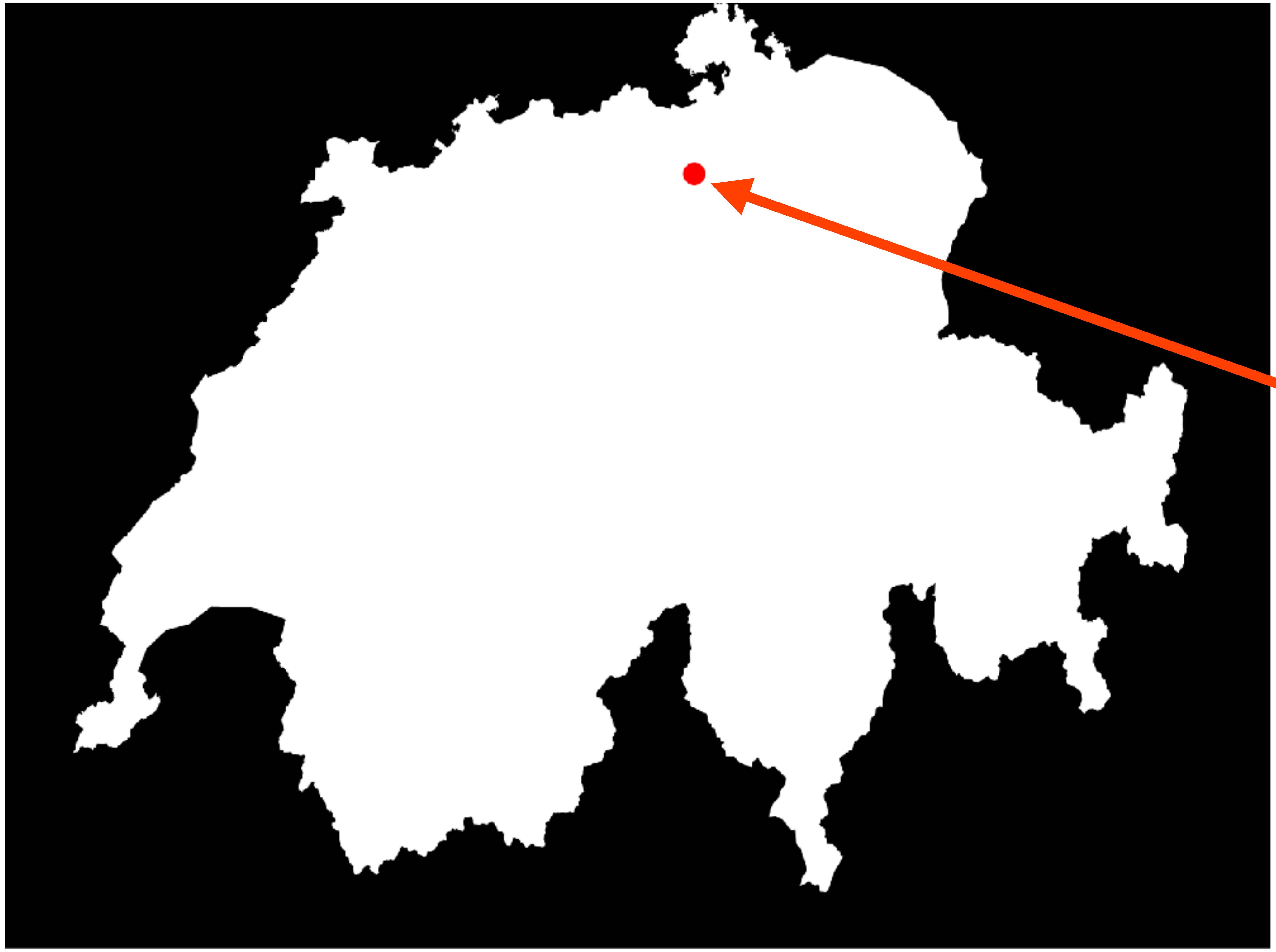
- **Geometry object:** This is either the point, line, or polygon described earlier. Basically the location information.

- **Feature object:** This is the geometry object and the associated random ad hoq data. Remember, GeoJSON doesn’t care what data you associate with the location information.

- **FeatureCollection:** Basically just a list of feature objects.

```
{
  "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "properties": {},
      "geometry": {
        "type": "Polygon",
        "coordinates": [
          [
            [
              7.200336456298827,
              45.87566836664006
            ],
            [
              7.2035980224609375,
              45.876863490676165
            ],
            [
              7.206344604492188,
              45.88009019718566
            ],
            [
              7.21527099609375,
              45.88379470321721
            ],
```

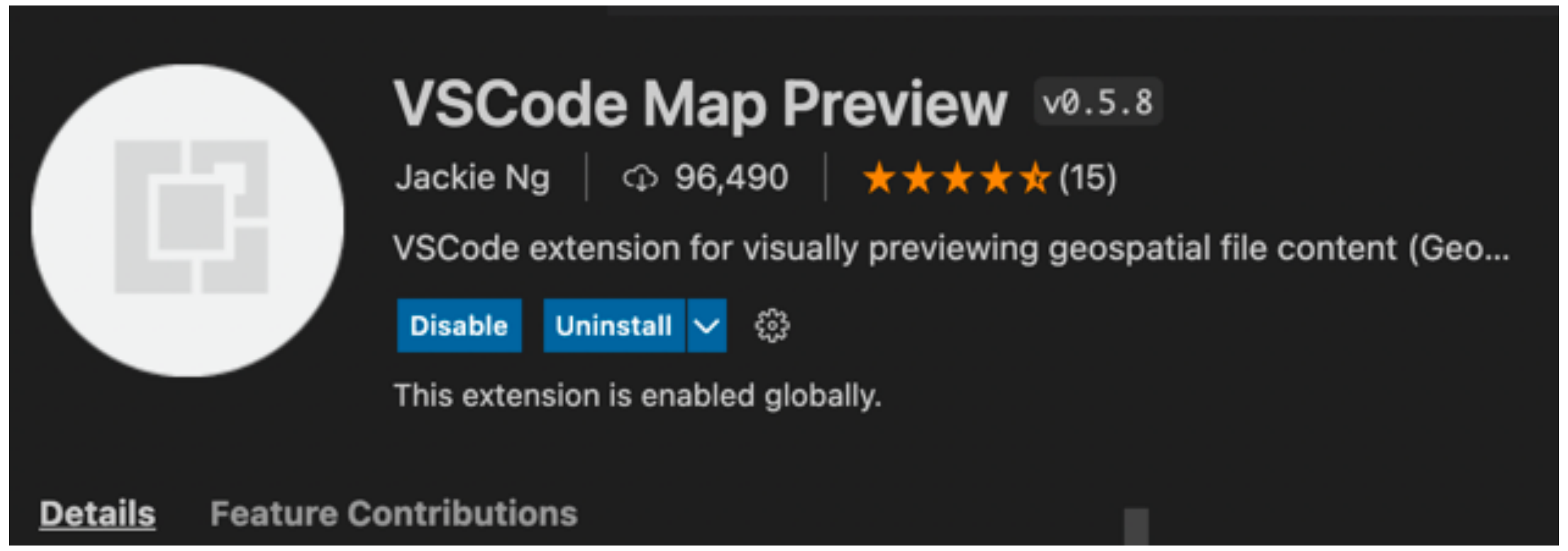
.geojson



Hey, this is Zurich!

.geojson

XYZ



VSCode Map Preview v0.5.8

Jackie Ng | 96,490 | ★★★★★ (15)

VSCode extension for visually previewing geospatial file content (Geo...

This extension is enabled globally.

.geotiff

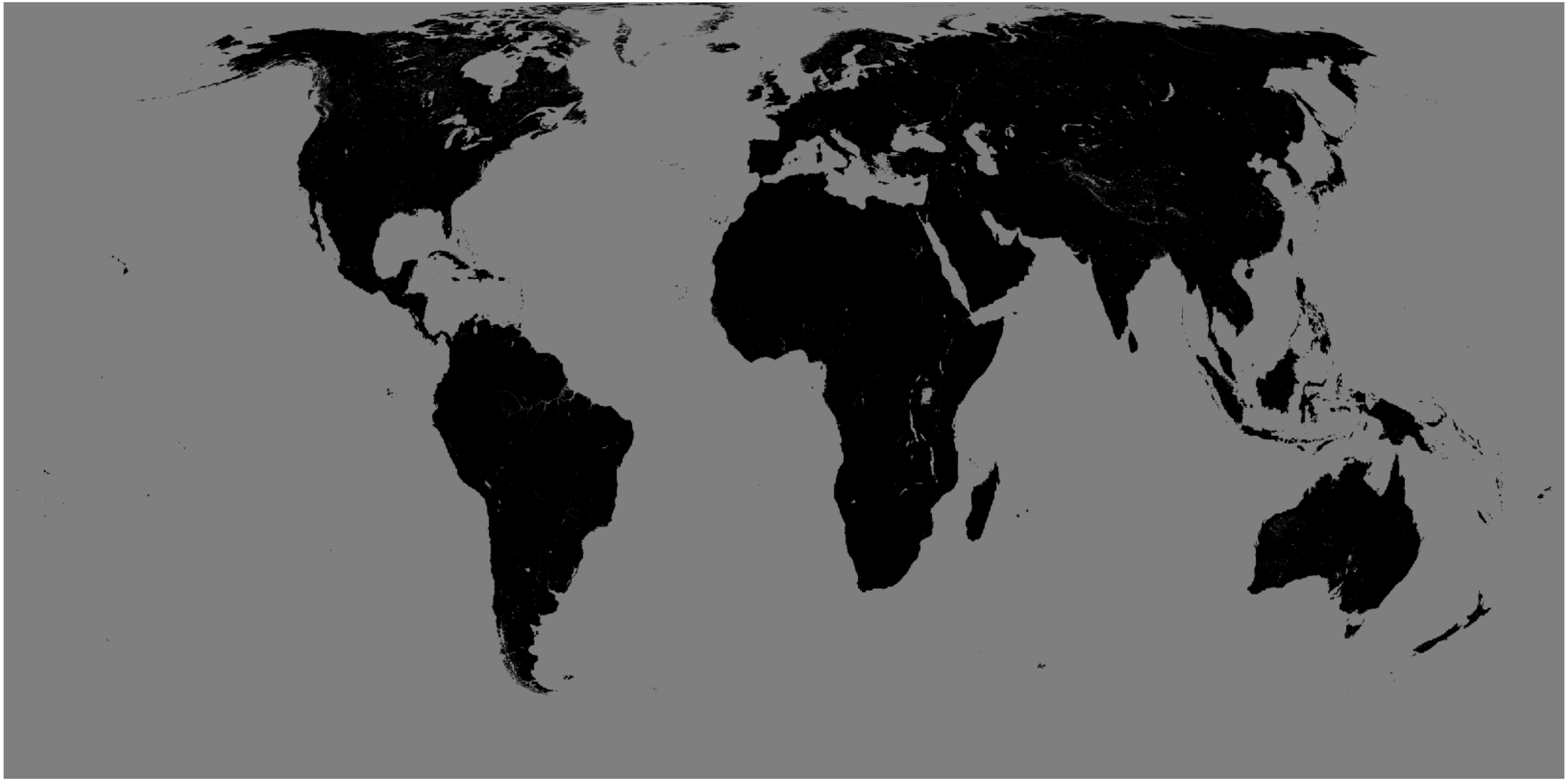
A **GeoTIFF** file extension contains geographic metadata that describes the actual location in space that each pixel in an image represents.

Spatial information is included in .tif file as embedded tags, which can include raster image metadata.

.geotiff

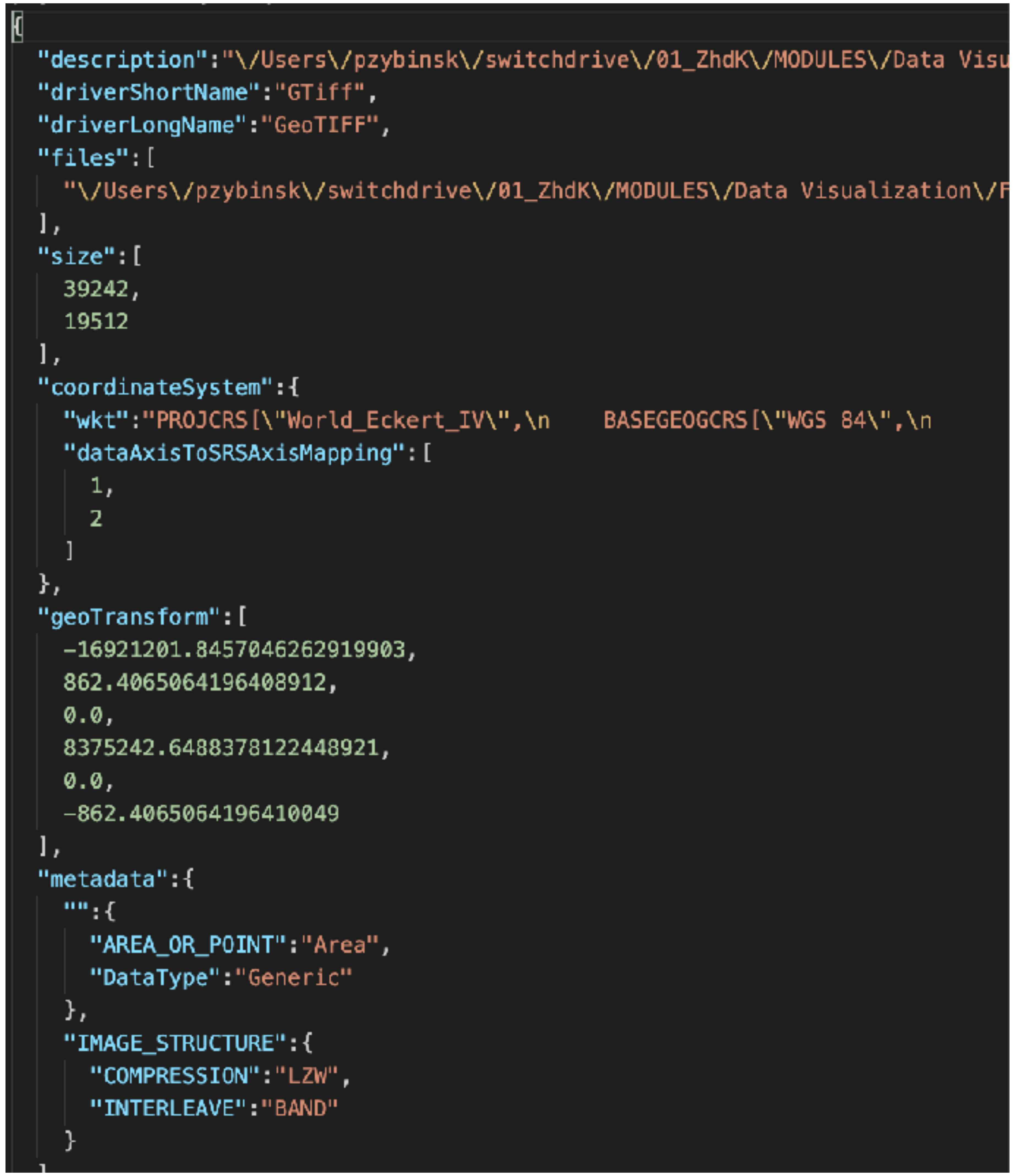
The **“Cloud Optimized GeoTIFF”** standard was developed in 2016 within the Open Source Geospatial Foundation/GDAL project as a format that is easily accessible with the use of common geographic software tools and web-based platforms.

.geotiff

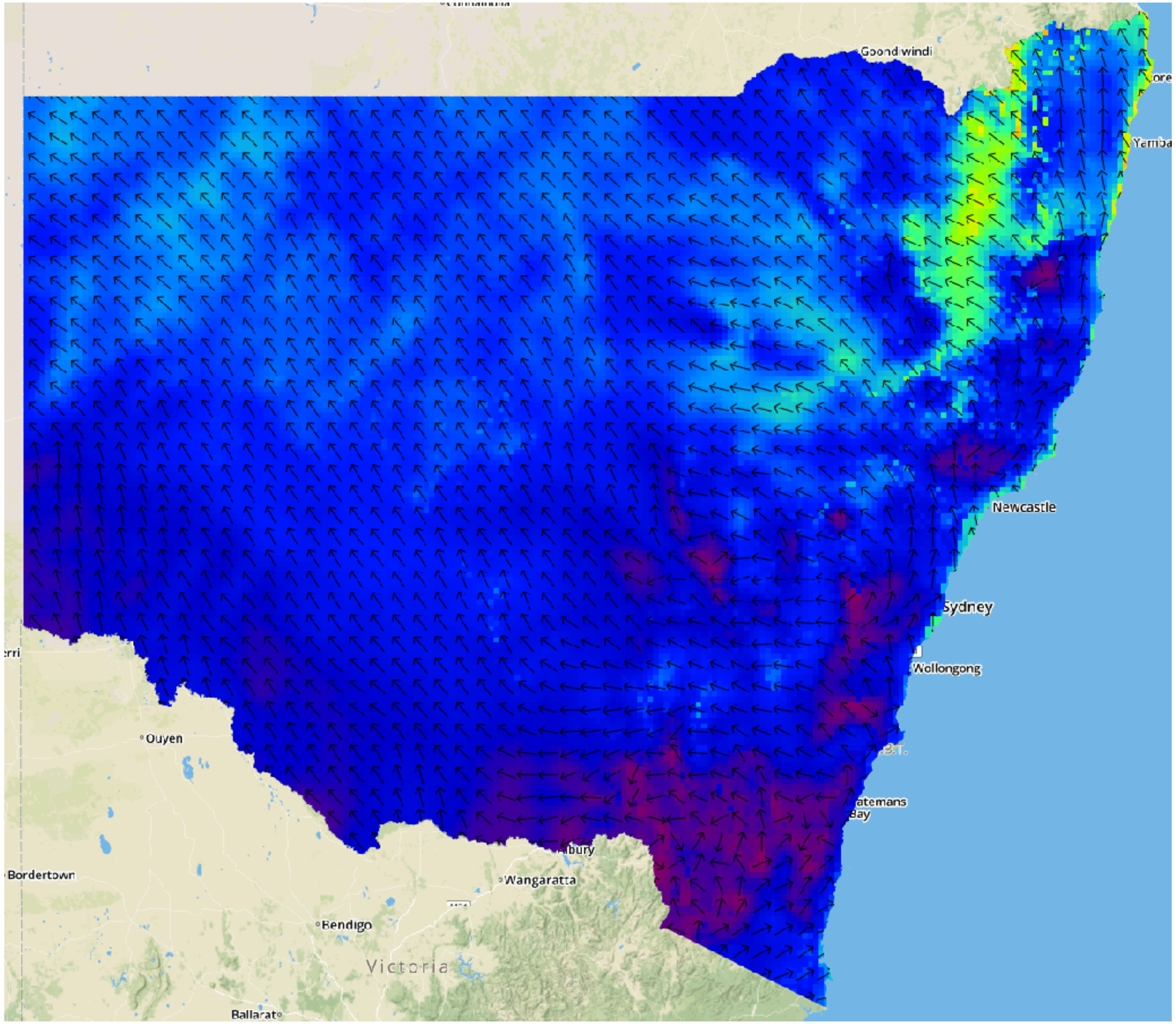


.geotiff

- horizontal and vertical datums
- spatial extent, i.e. the area that the dataset covers
- the coordinate reference system (CRS) used to store the data
- spatial resolution, measured in the number of independent pixel values per unit length
- the number of layers in the .tif file
- mathematical rules for map projection to transform data for a three-dimensional space into a two-dimensional display



```
{
  "description":"\/Users\/pzybinsk\/switchdrive\/01_ZhdK\/MODULES\/Data Visu
  "driverShortName":"GTiff",
  "driverLongName":"GeoTIFF",
  "files":[
    "\/Users\/pzybinsk\/switchdrive\/01_ZhdK\/MODULES\/Data Visualization\/F
  ],
  "size":[
    39242,
    19512
  ],
  "coordinateSystem":{
    "wkt":"PROJCRS[\"World_Eckert_IV\",\n    BASEGEOGCRS[\"WGS 84\",\n
    "dataAxisToSRSAxisMapping":[
      1,
      2
    ]
  },
  "geoTransform":[
    -16921201.8457046262919903,
    862.4065064196408912,
    0.0,
    8375242.6488378122448921,
    0.0,
    -862.4065064196410049
  ],
  "metadata":{
    "":{
      "AREA_OR_POINT":"Area",
      "DataType":"Generic"
    },
    "IMAGE_STRUCTURE":{
      "COMPRESSION":"LZW",
      "INTERLEAVE":"BAND"
    }
```


Goondiwindi
Yamba
Newcastle
Sydney
Wollongong
Batemans Bay
Albury
Wangaratta
Bendigo
Victoria
Ballarat
Ouyen
Bordertown

geotiff.js & geospatial

https://github.com/sacridini/Awesome-Geospatial#javascript

https://dataviz.cafe/

https://github.com/geotiffjs/geotiff.js

https://towardsdatascience.com/geotiff-coordinate-querying-with-javascript-5e6caaaf88cf