

eCFR Agency Rule Explorer – Project Summary

This project is a Python-based tool that extracts, organizes, and analyzes the structure of U.S. federal regulations from the Electronic Code of Federal Regulations (eCFR), with a focus on mapping regulatory rules to their responsible agencies.

DOGE eCFR Requirements:

- Please write code to download the current eCFR and analyze it for **items** such as **word count** per **agency** and historical changes over time. Feel free to add your own custom metrics.
-
- Build a complete mapping from government agencies to their associated CFR titles, chapters, subchapters, parts, and sections
- Provide detailed metrics for each agency's regulatory scope
- Handle the hierarchical and variable structure of the CFR, including optional subchapters
- Output clean, structured data for use in downstream tools like Power BI, Excel, or regulatory audit platforms

Post Analysis eCFR Requirements:

After reviewing the eCFR API, system requirements were refined to meet all original goals while incorporating bonus features and reducing data usage within the required timeframe. The enhanced requirements are as follows:

- Please write code to download the current eCFR Agency data.
- Once downloaded, analyze the data and provide the following word counts per Agency (Item):
 - Total Number of Titles per Agency
 - Total Number of Chapters per Agency per Titles
 - Total Number of Subchapters per Agency per Titles per Chapters
 - Total Number of Parts per Agency per Titles per Chapter per Subchapter (if present)
 - Total Number of Sections per Agency per Chapter per Subchapter (if present) per Part
 - Total Number of Historical changes historical changes over time.
 - ADDED METRICS

Key Objectives:

- Build a complete mapping from government agencies to their associated CFR titles, chapters, subchapters, parts, and sections
- Provide detailed metrics for each agency's regulatory scope
- Handle the hierarchical and variable structure of the CFR, including optional subchapters
- Output clean, structured data for use in downstream tools like Power BI, Excel, or regulatory audit platforms

How It Works

1. Agency Discovery

The script queries the official eCFR Admin API to retrieve a full list of federal agencies and their CFR references (title, chapter, or subtitle).

2. Title Versioning

For each title referenced by an agency, it fetches the **latest version** of the structure using the eCFR Versioner API.

3. CFR Structure Parsing

The script recursively walks through the CFR hierarchy:

- Title → Chapter → Subchapter (optional) → Part → Section

4. Grouping and Labeling

- All **Parts** are grouped under their parent **Subchapters** (or "No Subchapter" if none exists)
- **Sections** are listed under each Part
- Labels use the official eCFR format (e.g., "Part 60–Performance Standards")

5. Metrics Collection

For each agency, the script tracks:

- Number of Titles referenced
- Chapters and Subchapters found
- Parts and Sections counted

6. Output

The full data is written to a clean `out.json` file, which contains:

- The agency name
- CFR parts grouped by subchapter
- A full breakdown of section labels
- Summary metrics per agency

Output Example

```
{
  "agency": "Environmental Protection Agency",
  "parts_by_reference": [
    {
      "title": 40,
      "chapter": "Chapter I–Environmental Protection Agency",
      "parts_by_subchapter": [
        {
          "subchapter": "Subchapter C–Air Programs",
          "parts": [
            {
              "part": "Part 60–Standards of Performance",
              "sections": ["$ 60.1–Scope", "$ 60.2–Definitions"]
            }
          ]
        }
      ]
    }
  ],
  "metrics": {
    "title_count": 1,
```

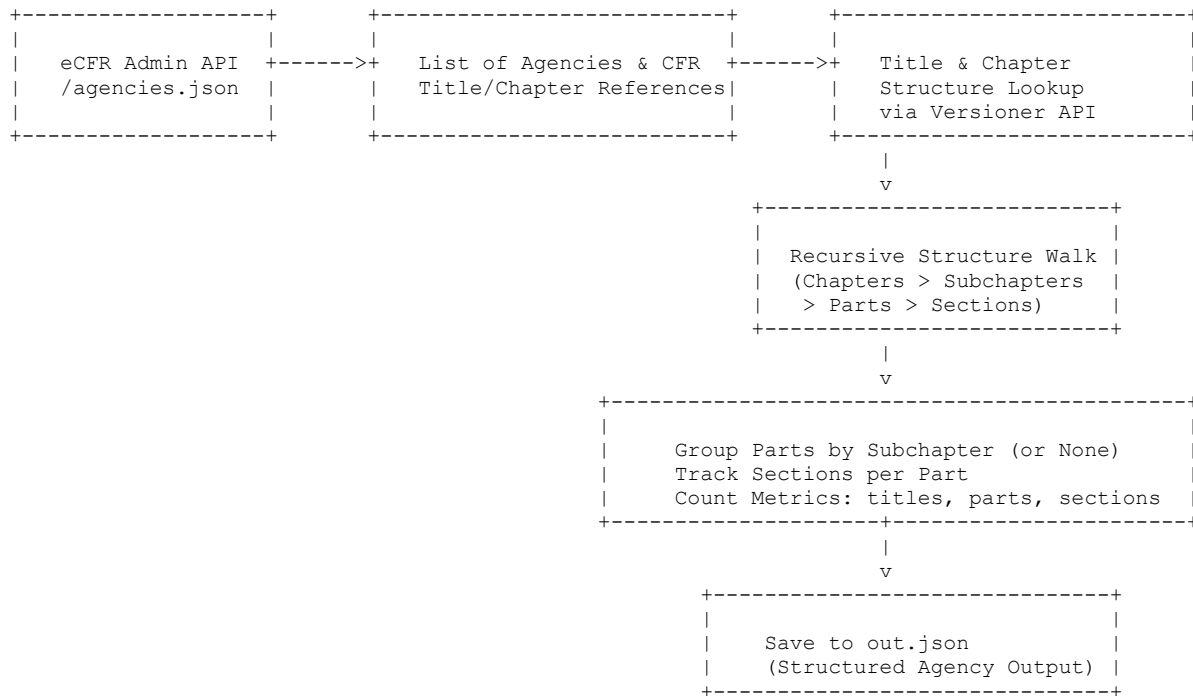
```
"chapter_count":      1,  
"subchapter_count":   1,  
"part_count":         20,  
"section_count":     310  
}  
}
```

Use Cases

- **Compliance Reporting**
Easily determine which parts of the CFR fall under a specific agency's domain.
- **Legal/Regulatory Analysis**
Support audits or reviews by surfacing the structure and scope of federal regulations.
- **Data Pipelines**
Integrate clean CFR structure into data warehouses or reporting systems (e.g., Power BI dashboards).
- **Civic Tech & Public Transparency**
Make federal rules more navigable and understandable for non-experts.

eCFR Agency Rule Explorer - Architecture Diagram

Data Flow Overview



Output Structure

- agency: Agency name (e.g., Environmental Protection Agency)
- parts_by_reference:
 - title, chapter
 - parts_by_subchapter
 - subchapter name
 - parts
 - part: Part label
 - sections: List of section labels
- metrics:
 - title_count, chapter_count, subchapter_count, part_count, section_count

Technologies Used

- Python 3
- Requests library
- JSON API from <https://www.ecfr.gov/developers/documentation/api/v1>

Output File

- out.json: Contains structured data for all agencies, suitable for auditing, analysis, and visualization.

Design Notes

- Designed to handle two types of CFR structures:
 - title > chapter > subchapter > part > section
 - title > chapter > part > section (no subchapter)
- Recursively traverses children nodes to dynamically discover structure.
- Uses `label` field to provide readable identifiers for parts/sections.