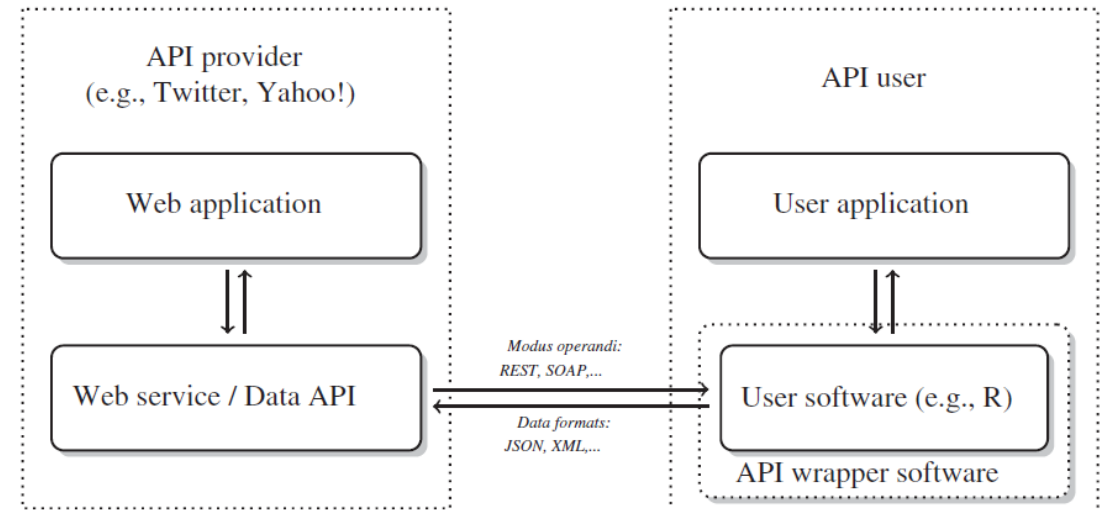# APIs

Akitaka Matsuo

Essex IADS

APIs

- API = Application Programming Interface
- Simple data transfer system for application
  - Provider grant access
  - Data are provided in a reliable, lightweight format

Munzert et al 2014, p. 259

Types of APIs

- **RESTful APIs**: Extract static information. Send the query to the server, receive the data (e.g. petition data)
- **Streaming APIs**: get the streaming of data generated in real-time. Once the connection to the server is established, new data keep flowing in (e.g. twitter streaming API)

To access an API

- Look for the documentation
  - Examples:
    - https://documents.worldbank.org/en/publication/documents-reports/api
    - https://developer.parliament.uk/
  - As the distribution of the data is the motivation, documentation is generally good
  - **Endpoints** and **parameters** are necessary to access
  - Rate-limits?
    - When the number of items is big, it's just impractical to get all.

Process of connecting with APIs

- Construct a URL for API call (if the method is GET)
  - Endpoint + parameters
- Request the data using the URL
- Get the response, parse
  - Common data formats: JSON, XML
- Some APIs require authentication
  - access key/token
  - OAuth

An example of RESTful API

- Parliament members
- Documentation:
  - https://members-api.parliament.uk/index.html
- Endpoint:
  - https://members-api.parliament.uk/api/Members/{id}
  - Parameters:
  - Members id
    - e.g 172
  - Output data format (.json, .xml)

API response (in JSON)

- URL:
  http://lda.data.parliament.uk/members/172.json
- Using urlopen
- The output will be handed to **json** package

{ "format" : "linked-data-api", "version" : "0.2", "result" : {"_about" :
"http://eldaddp.azurewebsites.net/members/172.json", "definition" :
"http://eldaddp.azurewebsites.net/meta/members/_id.json", "extendedMetadataVersion"
: "http://eldaddp.azurewebsites.net/members/172.json?_metadata=all", "primaryTopic"
: {"_about" : "http://data.parliament.uk/members/172", "additionalName" : {"_value"
: "Julie"}
    , "constituency" : {"_about" : "http://data.parliament.uk/resources/146966",
"label" : {"_value" : "Hackney North and Stoke Newington"}
        }
    , "familyName" : {"_value" : "Abbott"}
    , "fullName" : {"_value" : "Ms Diane Abbott"}
    , "gender" : {"_value" : "Female"}
    , "givenName" : {"_value" : "Diane "}
    , "homePage" : "http://www.dianeabbott.org.uk", "isPrimaryTopicOf" :
"http://eldaddp.azurewebsites.net/members/172.json", "label" : {"_value" :
"Biography information for Ms Diane Abbott"}
    , "party" : {"_value" : "Labour"}
    , "twitter" : {"_value" : "https://twitter.com/HackneyAbbott"}
    }
    , "type" : ["http://purl.org/linked-data/api/vocab#ItemEndpoint",
"http://purl.org/linked-data/api/vocab#Page"]}
}

API response (in XML)

- URL:
  http://lda.data.parliament.uk/members/172.xml
- Using urlopen
- The output will be handed to **xml** package

```xml
<?xml version="1.0" encoding="UTF-8"?>
<result format="linked-data-api"
        href="http://eldaddp.azurewebsites.net/members/172.xml"
        version="0.2">
  <definition href="http://eldaddp.azurewebsites.net/meta/members/_id.xml"/>
  <extendedMetadataVersion href="http://eldaddp.azurewebsites.net/members/172.xml?_metadata=all"/>
  <primaryTopic href="http://data.parliament.uk/members/172">
    <isPrimaryTopicOf href="http://eldaddp.azurewebsites.net/members/172.xml"/>
    <additionalName>Julie</additionalName>
    <constituency href="http://data.parliament.uk/resources/146966">
      <label>Hackney North and Stoke Newington</label>
    </constituency>
    <familyName>Abbott</familyName>
    <fullName>Ms Diane Abbott</fullName>
    <gender>Female</gender>
    <givenName>Diane </givenName>
    <homePage href="http://www.dianeabbott.org.uk"/>
    <label>Biography information for Ms Diane Abbott</label>
    <party>Labour</party>
    <twitter>https://twitter.com/HackneyAbbott</twitter>
  </primaryTopic>
  <type>
    <item href="http://purl.org/linked-data/api/vocab#ItemEndpoint"/>
    <item href="http://purl.org/linked-data/api/vocab#Page"/>
  </type>
</result>
```

Compare API responses

University of Essex

JSON

```
{ "format" : "linked-data-api", "version" : "0.2", "result" : {"_about" :
"http://eldaddp.azurewebsites.net/members/172.json", "definition" :
"http://eldaddp.azurewebsites.net/meta/members/_id.json", "extendedMetadataVersion"
: "http://eldaddp.azurewebsites.net/members/172.json?_metadata=all", "primaryTopic"
: {"_about" : "http://data.parliament.uk/members/172", "additionalName" : {"_value"
: "Julie"}
        , "constituency" : {"_about" : "http://data.parliament.uk/resources/146966",
"label" : {"_value" : "Hackney North and Stoke Newington"}
        }
        , "familyName" : {"_value" : "Abbott"}
        , "fullName" : {"_value" : "Ms Diane Abbott"}
        , "gender" : {"_value" : "Female"}
        , "givenName" : {"_value" : "Diane "}
        , "homePage" : "http://www.dianeabbott.org.uk", "isPrimaryTopicOf" :
"http://eldaddp.azurewebsites.net/members/172.json", "label" : {"_value" :
"Biography information for Ms Diane Abbott"}
        , "party" : {"_value" : "Labour"}
        , "twitter" : {"_value" : "https://twitter.com/HackneyAbbott"}
    }
    , "type" : ["http://purl.org/linked-data/api/vocab#ItemEndpoint",
"http://purl.org/linked-data/api/vocab#Page"]}
}
```

XML

```
<?xml version="1.0" encoding="UTF-8"?>
<result format="linked-data-api"
        href="http://eldaddp.azurewebsites.net/members/172.xml"
        version="0.2">
    <definition href="http://eldaddp.azurewebsites.net/meta/members/_id.xml"/>
    <extendedMetadataVersion href="http://eldaddp.azurewebsites.net/members/172.xml?_metadata=all"/>
    <primaryTopic href="http://data.parliament.uk/members/172">
        <isPrimaryTopicOf href="http://eldaddp.azurewebsites.net/members/172.xml"/>
        <additionalName>Julie</additionalName>
        <constituency href="http://data.parliament.uk/resources/146966">
            <label>Hackney North and Stoke Newington</label>
        </constituency>
        <familyName>Abbott</familyName>
        <fullName>Ms Diane Abbott</fullName>
        <gender>Female</gender>
        <givenName>Diane </givenName>
        <homePage href="http://www.dianeabbott.org.uk"/>
        <label>Biography information for Ms Diane Abbott</label>
        <party>Labour</party>
        <twitter>https://twitter.com/HackneyAbbott</twitter>
    </primaryTopic>
    <type>
        <item href="http://purl.org/linked-data/api/vocab#ItemEndpoint"/>
        <item href="http://purl.org/linked-data/api/vocab#Page"/>
    </type>
</result>
```

API: Pros and Cons

- **Pros**
  - Data are well formatted
    - Parsing is trivial compared to web-scraping
  - Less frequently changed
    - Reliable, reproducible
  - API packages in Python/R for major ones
- **Cons**
  - Not all web-data are on APIs
  - Subject to the providers mercy
    - Changes, discontinuation