Dimensions

The Dimensions APIs

Harness the full power
of Dimensions data

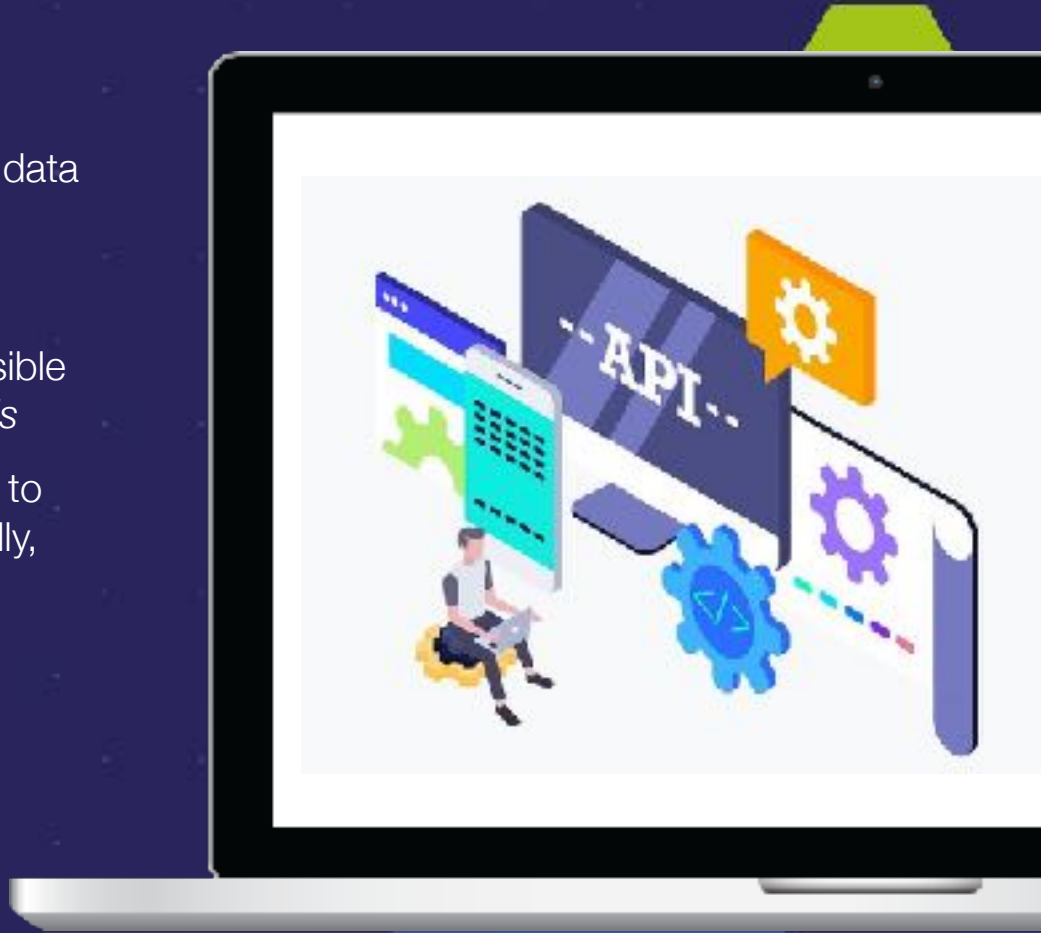
Michele Pasin
Head of Dimensions Data Solutions
m.pasin@digital-science.com

Part of DIGITAL science

What are Dimensions APIs?

APIs permit to carry out sophisticated research data analysis tasks. Common use cases are:

1. Queries which would be difficult or impossible to do via the web-app e.g. *citation analysis*
2. Automate query composition e.g. in order to build *reports* which are updated periodically, or *dashboards* with interactive controls
3. Get more data than what the web export functionality allows (50k results)



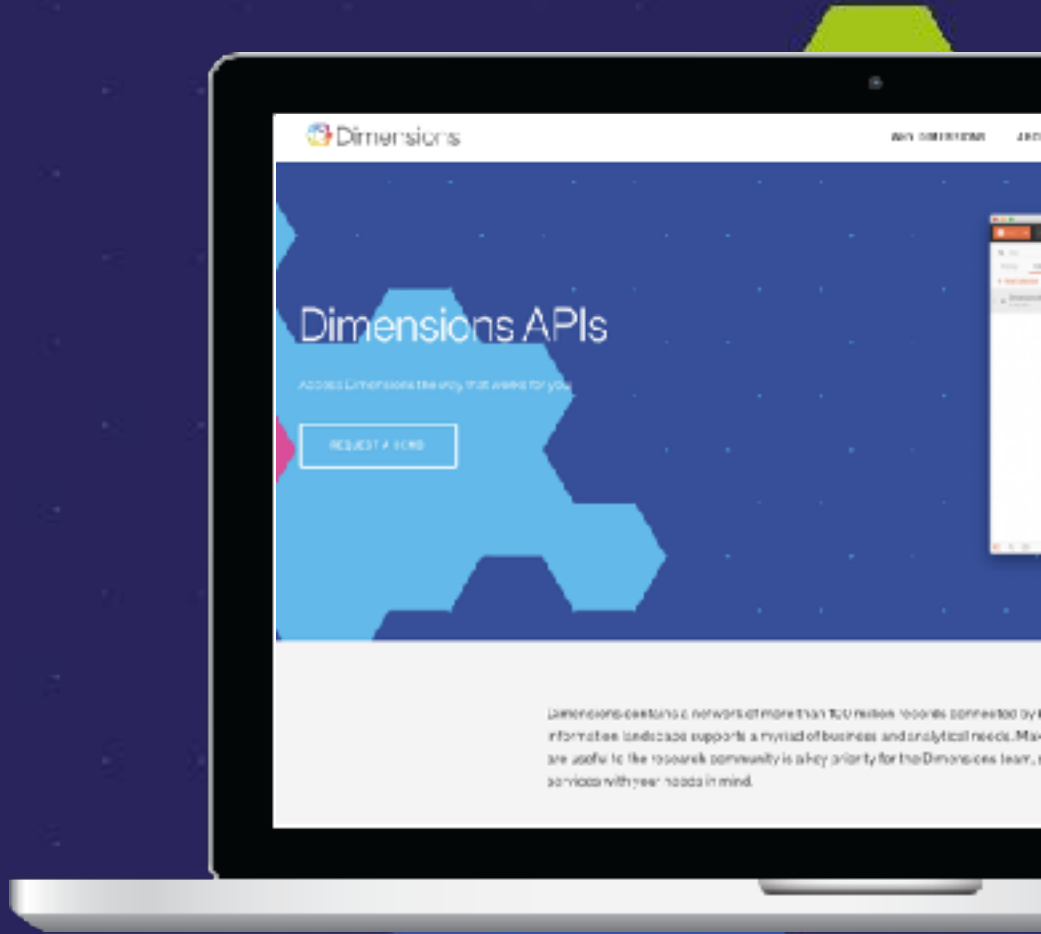
Dimensions APIs Products

Analytics API

- Cutting-edge features; frequent updates
- Throttled at 30 queries per minute
- Use case: analysts, data scientists

CRIS API

- Focus on extra stability and reliability; less frequent updates
- Throttled at 30 queries per minute
- Use case: integration with CRIS/RIMS



https://www.dimensions.ai/dimensions-apis/

Dimensions

APIs: choose your search paradigm

- **Full-text** search by keyword in millions of documents indexed by Dimensions
  - Stemming, relevancy scores, boolean operators, +/- syntax, proximity operations
  - Multiple indexes: *Title+abstract*, *Full data*, *Authors* etc..
- **Structured** metadata searches
  - Identifiers, Dates
  - Authors, Affiliations, References
  - Subject areas, Concepts



Dimensions Search Language (DSL)

Powerful query language developed around simple syntax, allowing users with various levels of technical skills to use.

- Basic query based around two phrases
- More complexity can be included by adding filters
- Returned data can be specified using custom fieldsets



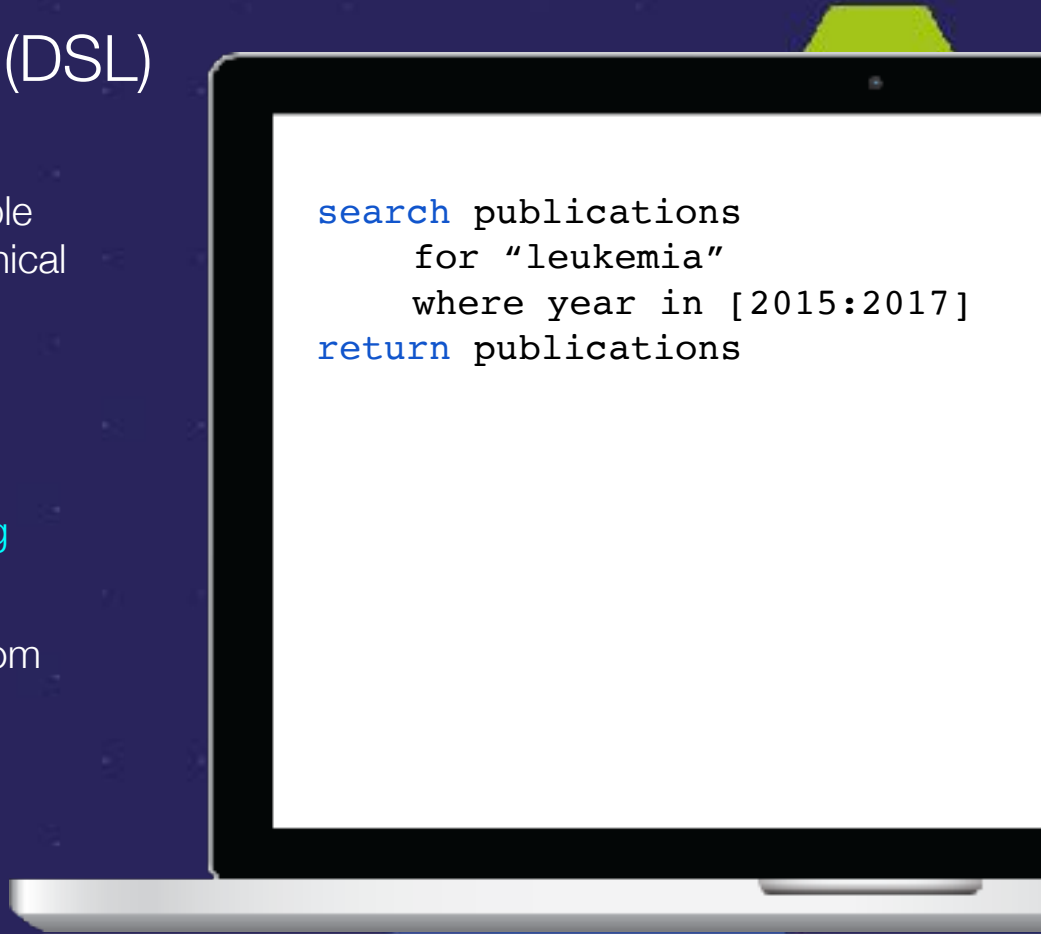
https://docs.dimensions.ai/dsl/language.html

Dimensions

Dimensions Search Language (DSL)

Powerful query language developed around simple syntax, allowing users with various levels of technical skills to use.

- Basic query based around two phrases
- More complexity can be included by adding filters
- Returned data can be specified using custom fieldsets



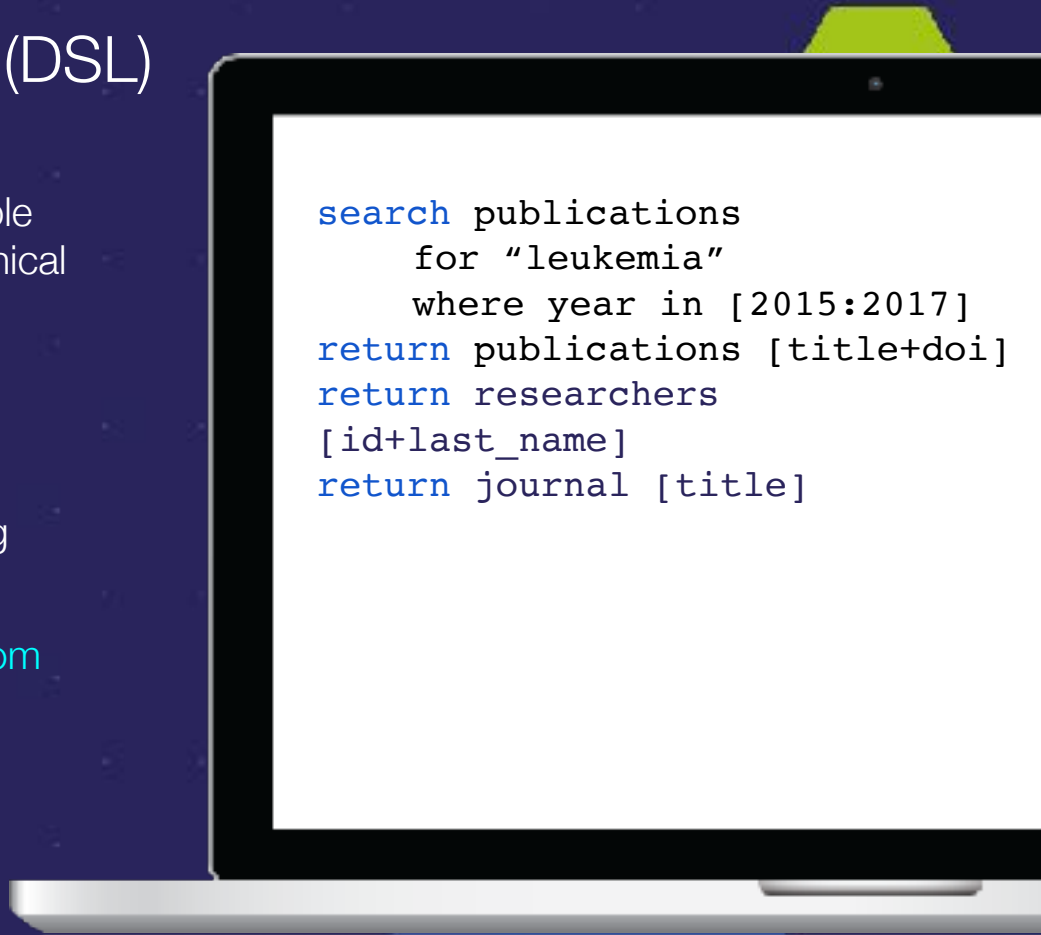
https://docs.dimensions.ai/dsl/language.html

Dimensions

Dimensions Search Language (DSL)

Powerful query language developed around simple syntax, allowing users with various levels of technical skills to use.

- Basic query based around two phrases
- More complexity can be included by adding filters
- Returned data can be specified using custom fieldsets



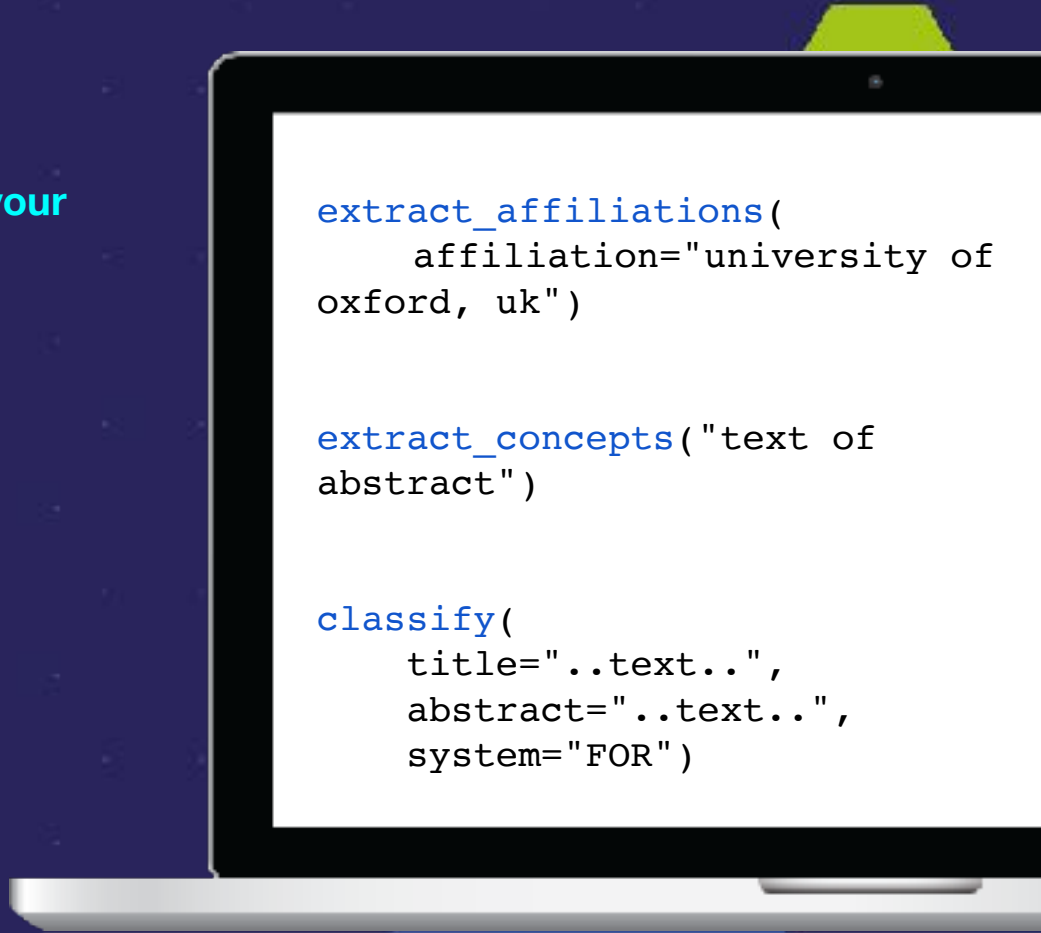
https://docs.dimensions.ai/dsl/language.html

Dimensions

DSL Special Functions

Make it easier to enrich / disambiguate your data

- **Affiliations** matching
  - From strings to GRID IDs
- **Concepts** extraction
  - From text to keywords
- **Classification** service
  - From text to ontologies



https://docs.dimensions.ai/dsl/functions.html

Dimensions

DSL Special Functions #2

Expert Identification workflow: permits to match researchers starting from text.
Applications: expert finding, reviewers selection

- **Two steps:** first Concept Extraction, then Expert Identification
- Using **publications** or **grants** sources, incl. arbitrary filters
- **Annotations**: Organizational conflict and Coauthorship conflict

```
identify experts

    from concepts "malaria
        OR \"effective malaria
        vaccine\"
        OR \"effective
        prevention\""

    using publications
        where year >= 2015

return experts[basics]
```

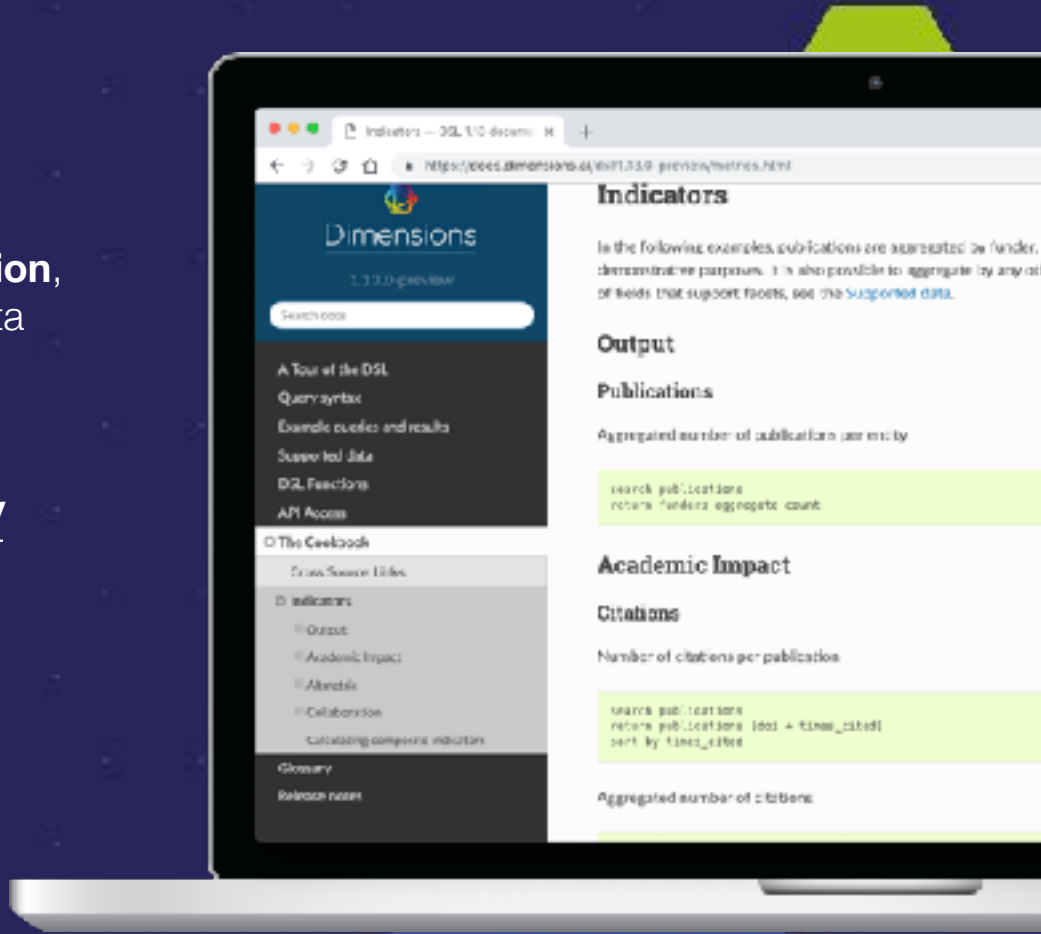
https://docs.dimensions.ai/dsl/expert-identification.html

Dimensions

Online Documentation

The Dimensions API has extensive **documentation**, containing examples, tutorials and supported data tables

https://docs.dimensions.ai/dsl/latest/



API LAB: reusable Notebooks

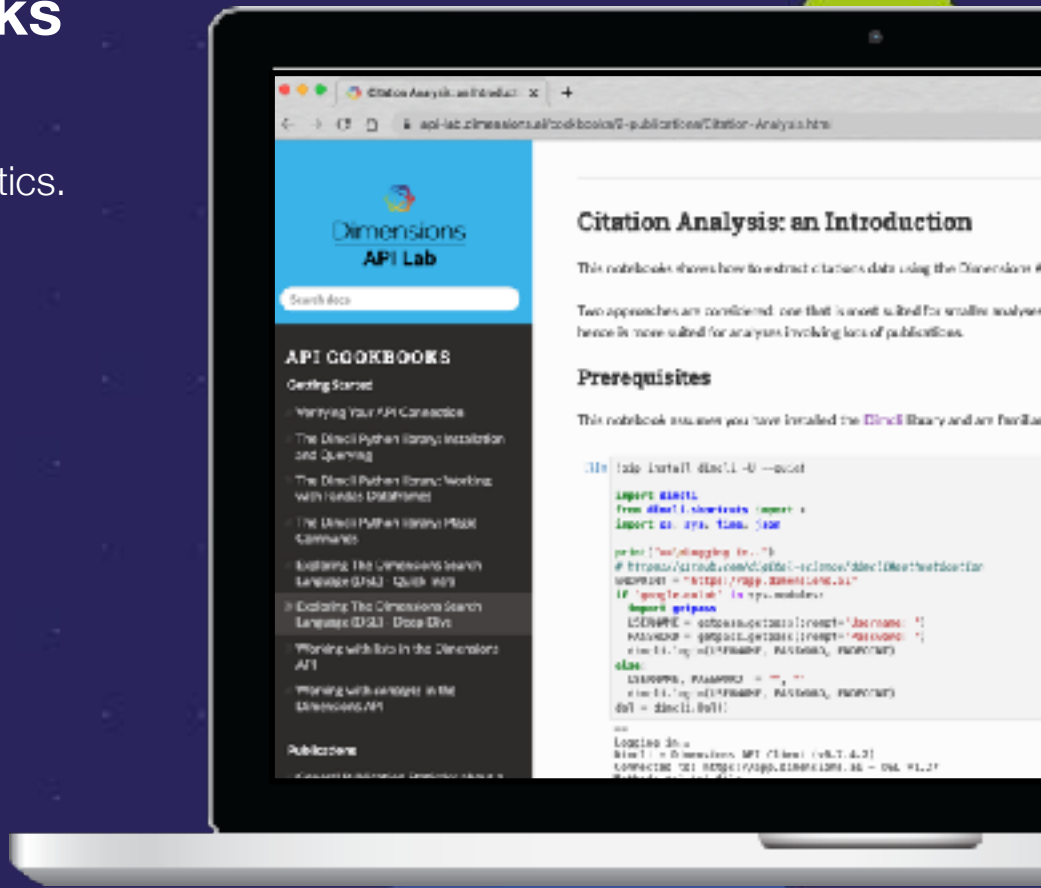
Open source collection of ready-to-use Jupyter Notebooks using the API for scholarly data analytics.

- Fully compatible with *Google Colab*!

For example:

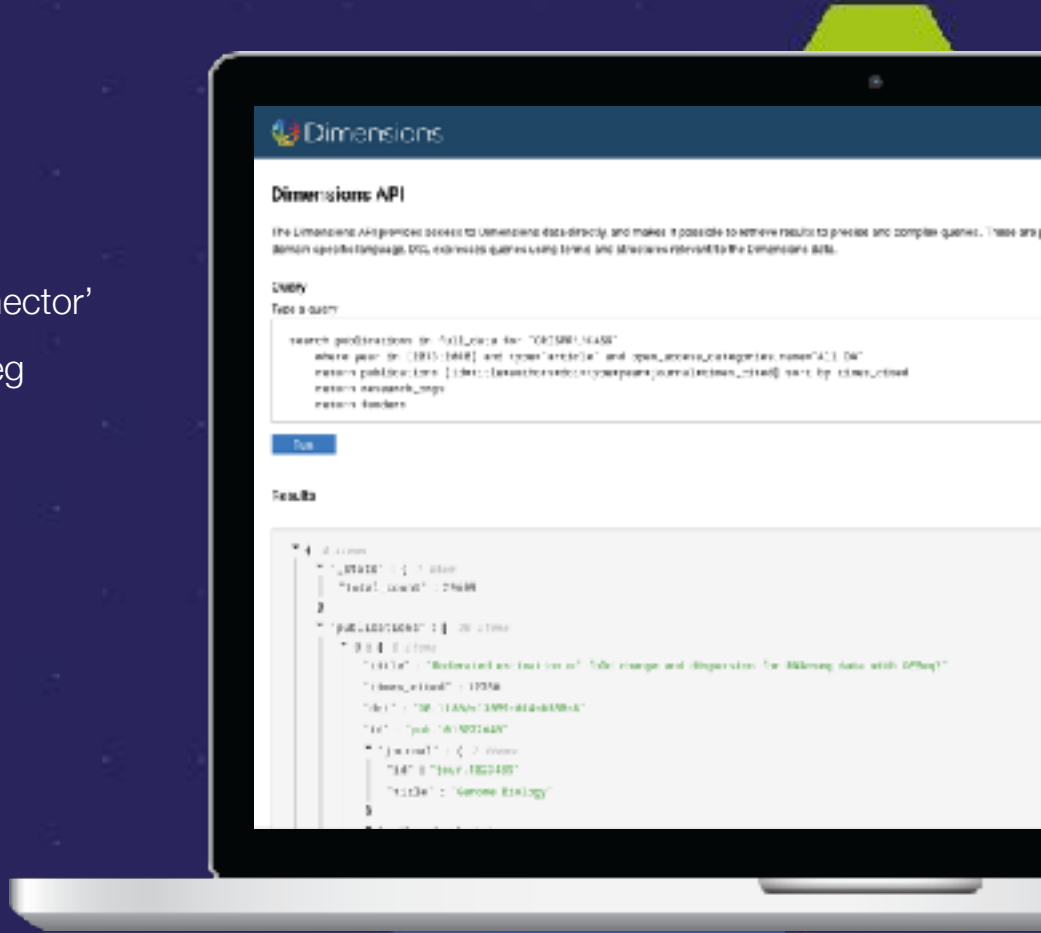
- Calculating the H Index of a researcher
- International and Industry collaboration
- Building a Citation Graph
- Extracting Patents Citing an Organization

https://api-lab.dimensions.ai/



How to Access the API

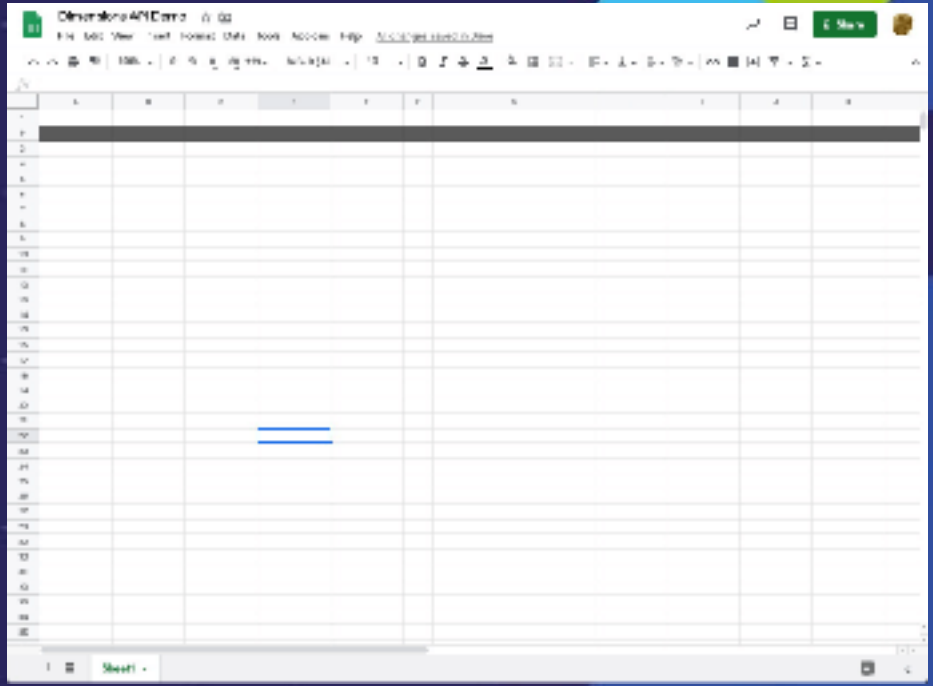
- Via the Dimensions app (link)
  - Handy for testing purposes
- Via third-party softwares that have a ‘data connector’
  - Simplifies integration with existing tools, eg Qlik, Tableau, Google Sheets..
- Via a programming language
  - Ad-hoc analysis and data processing, eg
    - Python: *Dimcli* library
    - R: *dimensionsR* library
    - R: *rDimensions* library



Gsheets Data Connector

- Query the API from Google Sheets
- Automatically turn JSON query results into spreadsheets
- Analyse the data using gsheet charts, pivot tables etc..

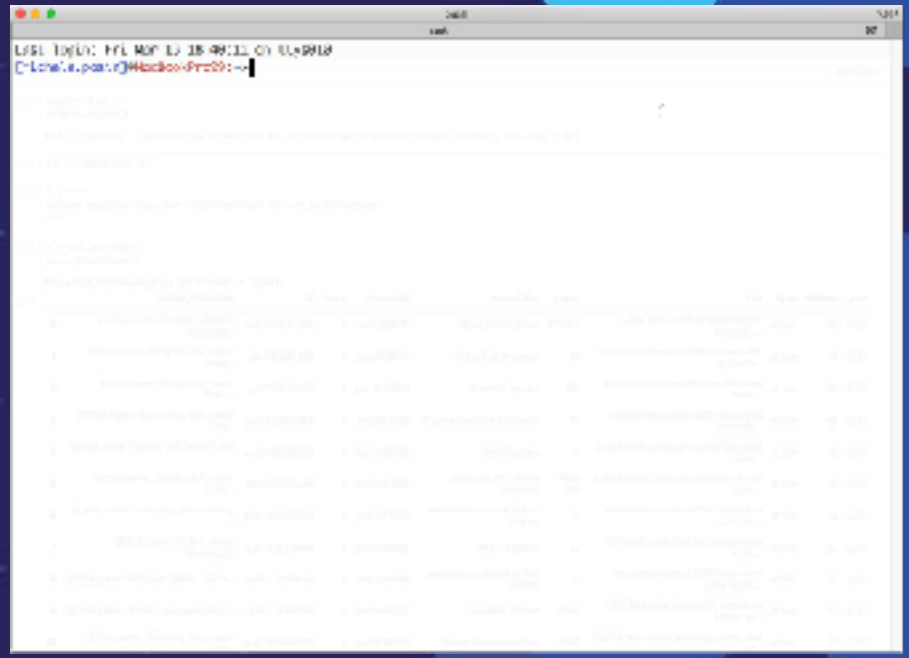
Available in the Gsuite marketplace



https://gsuite.google.com/marketplace/app/dimensions_api/961965218337

DIMCLI: API Python Client

- Simplifies common API operations
  - Log in, log out
  - Querying, iterative queries
  - Extracting and transforming data e.g. to dataframes
- Includes a Command Line Interface (CLI)
  - Like a ‘query console’ with autocomplete and other handy functionalities



https://github.com/digital-science/dimcli

Today's hands-on session

Analysing NIH funded researchers

Outline

1. We get all NIH Dimensions identifiers
2. We explore NIH grants based on chosen set of criteria eg years or keywords, in order to quickly discover trends
3. We analyse who else is funding the NIH-funded researchers



https://github.com/digital-science/dimensions-api-lab/blob/master/archive/2021-11-10-NIH-Training/2021-11-NIH-funding-history.ipynb

Dimensions

Dimensions

Thanks!

Email

supportapi@dimensions.ai

Users forum: https://groups.google.com/forum/#!forum/dimensions-api-users

Part of DIGITAL science