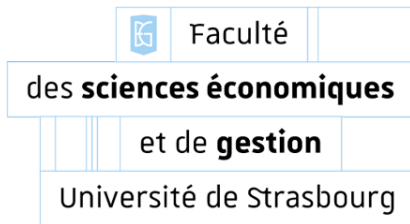# Novelty Components of Scientific Productions
-
Data Visualization

Supervised by: Diletta Abbonato
Authors: Valentin Barthel, Adrien Busché, Diego Fracassi

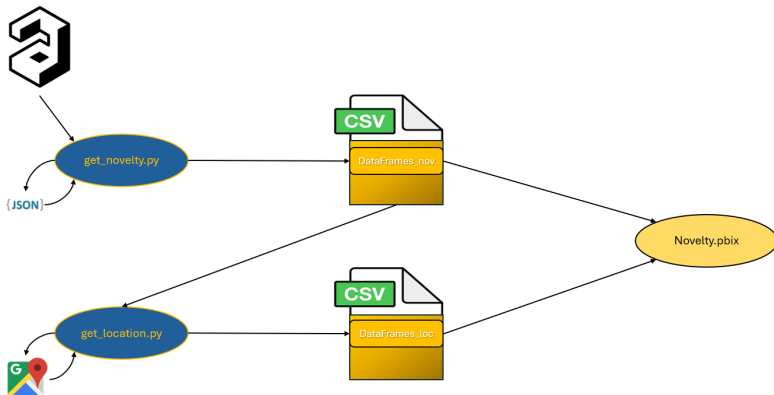valentin.barthel@etu.unistra.fr
adrien.busche@etu.unistra.fr
diego.fracassi@etu.unistra.fr

Table des Matières

Introduction

Our Program Diagram

$\implies Github$

Co-occurrence Matrix

- The co-occurrence matrix represents how frequently two references (journals) appear together in a scientific article.
- Each cell (i, j) contains the number of times journal i and journal j are cited together.

Reference	Journal A	Journal B	Journal C	Journal D
Journal A	10	3	2	0
Journal B	3	15	1	1
Journal C	2	1	12	0
Journal D	0	1	0	8

Table: Example of a co-occurrence matrix

$\implies get\_novelty.py$

Geocoding

$\implies get\_location.py$

Limitations of the Novelty Indicator

- **Dependence on citation patterns:** The indicator relies heavily on the availability and accuracy of citation data.
- **Field-specific dynamics:** It may not account for differences in citation practices across disciplines.
- **Temporal lag:** It only reflects novelty after a reference has been cited, potentially overlooking emerging ideas not yet widely cited.
- **Computational limitations:** In very large datasets, constructing co-occurrence matrices and calculating frequencies can be resource-intensive.
- **Simplistic measure of novelty:** The indicator may not capture the full spectrum of creativity or innovation in the research being analyzed.