

(Career) benefits of open data



*Antonio Schettino, Ph.D.
Coordinator Open Science
Erasmus Research Services
schettino@eur.nl*



OUTLINE



- What is Data?
- What is Open Data?
- Benefits of Open Data
 - for individual careers
 - for research discipline
 - for society

The Erasmus logo, featuring a stylized, handwritten-style script of the word "Erasmus" in black.

Antonio Schettino



WHAT IS DATA?

Erasmus

Antonio Schettino



WHAT IS DATA?



“A representation of facts, concepts or instructions in a formalised manner suitable for communication, interpretation, or processing by humans or by automatic means” (Hicks, 1993)

Data can *include*:

- Spreadsheets
- Text extracts of images
- Maps
- Archival documents
- Photographs, audio, film recordings
- Field notes
- Interview notes
- ...

Antonio Schettino



WHAT IS OPEN DATA?

Erasmus

Antonio Schettino



OPEN SCIENCE



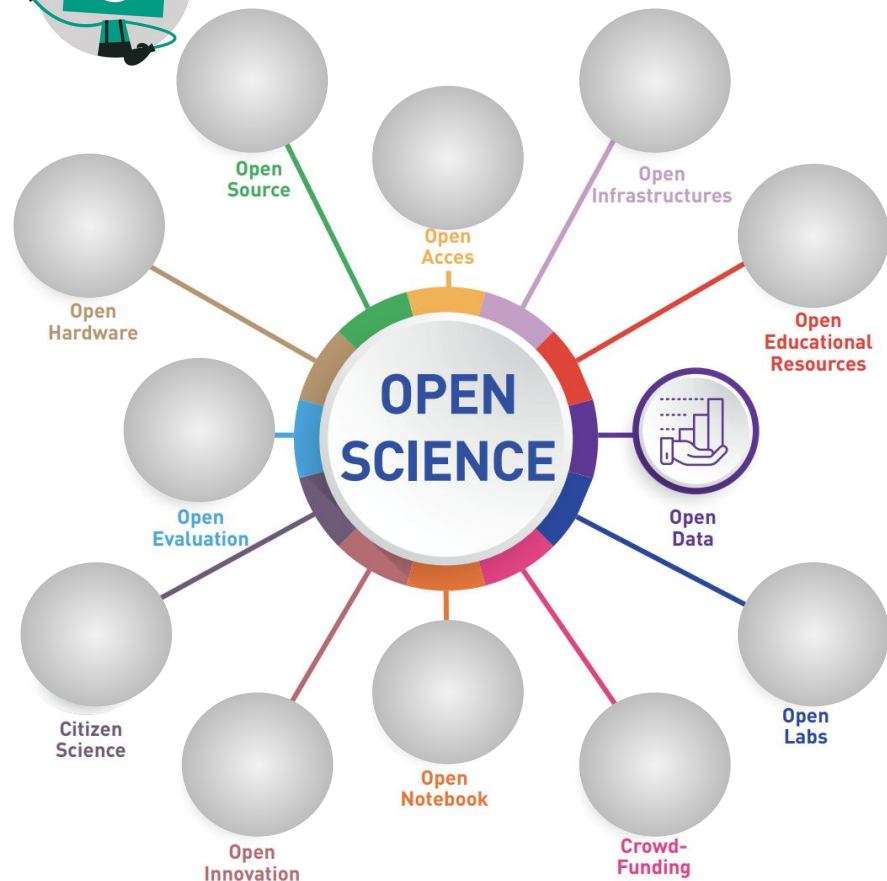
Erasmus

Towards a UNESCO Recommendation on Open Science

Antonio Schettino



OPEN DATA



- online publication of data
- available for access, (re-)use, (re-)distribution
- permissive licenses (e.g., *Creative Commons*)
- human- and machine-readable (e.g., *FAIR*)
- example:
 - *COVID-19 in EU/EEA (ECDC)*
 - *DANS Easy*

Erasmus



BENEFITS OF OPEN DATA: INDIVIDUAL CAREERS

Erasmus

Antonio Schettino



REQUIRED BY FUNDERS



- **NWO**

- *“As a **minimum**, share the research data that underlie research publications alongside those publications, unless this is prevented for reasons of privacy, public safety, ethical restrictions, property rights or commercial interests”*

- **Horizon Europe [2021-2027]**

- *“projects generating research data **must** manage their data responsibly and in line with **FAIR principles**” (no opting out)*
- *open access to research data “as open as possible, as closed as necessary”*

- **Bill & Melinda Gates foundation**

- *“unrestricted access and reuse of all Funded Developments [...] including any underlying data sets”*



CITATION ADVANTAGE



- studies with available data receive **5% - 13% more citations** than similar studies with no data
- major contributor: **data reuse**, either by original authors or other researchers
- citation of **up to 40%** in *American Economic Review* and *American Journal of Political Science*, where policies are **enforced**
- **qualitative** datasets are often used for *bachelor/master theses and teaching*

Antonio Schettino



JOB COMPETITIVENESS



Academic job offers that mentioned open science

Contributors: Felix Schönbrodt, David Thomas Mellor, Christina Bergmann, Naomi Penfold, Samuel Westwood, Alexandra Lautarescu, Olivia Kowalczyk, Lorenza Dall'Aglio, Elisabet Blok, Antonio Schettino,

Leonhard Falk Florentin Schramm, Bianca Weber, Greg Tananbaum

Date created: 2018-01-18 02:25 PM | Last Updated: 2021-07-22 01:27 PM

Identifier: DOI 10.17605/OSF.IO/7JBNT

Category: Project

Description:

This project archives job offers (mostly professorship positions) which require or suggest an open science statement from the applicants.

Wiki

FOR CONTRIBUTORS

Recommended workflow:

1. name uploaded files as `YEAR_COUNTRY_CITY_POSITION_TOPIC` (example: `2020_UK_Oxford_PostDoc_CognitivePsychology.pdf`) into the correct folder.
2. Also add the meta-information into the [Google Sheet database](#).

Please don't forget the second step, otherwise the PDFs and the meta-data sheet get out of sync, which is tedious to fix.

Quick and dirty workflow, for the bu...

[Read More](#)

Files

As of October 2021:

- 142 vacancies (since 2017)
- 7 countries (including NL)
- <https://osf.io/7jbnt/>

Antonio Schettino



BENEFITS OF OPEN DATA: RESEARCH DISCIPLINE

Erasmus

Antonio Schettino



CREDIBILITY



- Trustworthiness and confidence
 - *High-powered replications of laboratory experiments in economics*
 - $p < .05$: 100% in original studies, **61% in replications**
 - mean effect size (r): .47 in original studies, **.28 in replications**
 - *Uncommon replication research in management and business*
 - **< 10%** of all published papers can be classified as replication
 - **48%** provided **no support** for original findings
- Fraud detection
 - *Dirk Smeesters*: violation of scientific integrity
 - *"The most striking problem in investigating the scientific integrity of the papers by Smeesters was the difficulty in obtaining the data"*



BENEFITS OF OPEN DATA: SOCIETY

Erasmus

Antonio Schettino



FINANCIAL CRISIS AND AUSTERITY MEASURES



Growth in a Time of Debt (Carmen Reinhart & Kenneth Rogoff)

- debt > 90% gross domestic product → drop in economic growth
- “justify” switch from stimulus to austerity (example: Greece)
- **However:** peers’ criticism from the start
 - unwarranted claims of causality
 - other data did not show the same trend
- when Reinhart & Rogoff shared their spreadsheet:
 - some data omitted in calculations due to coding error
 - unconventional weighting of summary statistics
 - correct analysis - and *recent data* - provide **no evidence** of low GDP growth when countries hit 90% debt-to-GDP ratio

The New Yorker (2013)

Antonio Schettino



SARS-CoV-2



GenBank ▾

Send to: ▾

Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome

GenBank: MN908947.3

[FASTA](#) [Graphics](#)

Go to: ☒

LOCUS MN908947 29903 bp ss-RNA linear VRL 23-JAN-2020
DEFINITION Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome.
ACCESSION MN908947
VERSION MN908947.3
KEYWORDS .
SOURCE Wuhan seafood market pneumonia virus
ORGANISM [Wuhan seafood market pneumonia virus](#)
Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; unclassified Betacoronavirus.
REFERENCE 1 (bases 1 to 29903)
AUTHORS Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Hu, Y., Song, Z.-G., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E.C. and Zhang, Y.-Z.
TITLE A novel coronavirus associated with a respiratory disease in Wuhan of Hubei province, China
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 29903)
AUTHORS Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Hu, Y., Song, Z.-G., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E.C. and Zhang, Y.-Z.

GenBank

Jan. 17th, 2020

Antonio Schettino



TIPS



- **plan, plan, plan!**
 - think of open data *before* starting your project
- **collaboration & cooperation**
 - share workload with colleagues
- **ask a professional (data stewards, library, ERS)**
 - advice, guidance, practical help

A stylized, handwritten signature in black ink.

Antonio Schettino



THANKS FOR YOUR ATTENTION!



Open Science Community Rotterdam

website: openscience-rotterdam.com

Twitter: [@OSCRotterdam](https://twitter.com/OSCRotterdam)

A handwritten signature in black ink, reading "Antonio Schettino".

Antonio Schettino