

# **Open science data** **(Dados abertos em ciência)**

**Marcos Duarte**

marcos.duarte@ufabc.edu.br

<http://demotu.org>

**Universidade Federal do ABC**

<https://git.io/JgfrO>

# Open science data

- “**Open science data** is a type of open data focused on publishing observations and results of scientific activities available for anyone to analyze and reuse.”
- Open science data maximizes scientific impact, reproducibility, reuse, and cost reduction.
- To open: data, software code ...

[https://en.wikipedia.org/wiki/Open\\_science\\_data](https://en.wikipedia.org/wiki/Open_science_data)

<https://gigascience.biomedcentral.com/articles/10.1186/s13742-016-0127-4>

# Main feature of Open Science Data

- Valuable.

...just like with anything else in science.

# Why Open Science Data matters to you

- You want to produce 'science' but don't have access to a laboratory with the necessary infrastructure;
- You don't have time to produce data to test your ideas;
- You need to learn, train or test something and need some data;
- You just want to help;
- ...

# Data Repositories

Generalists:

- [Figshare](#)
- [Open Science Framework](#)
- [Zenodo](#)
- [Mendeley Data](#)
- [Science Data Bank](#)

See full list at: <https://www.nature.com/sdata/policies/repositories>

# **A data paper (data descriptor)**

- A publication that provides detailed descriptions of research datasets, including the methods used to collect the data and technical analyses supporting the quality of the measurements.
- Data Descriptors focus on helping others reuse data, rather than testing hypotheses, or presenting new interpretations, methods or in-depth analyses.

# Journals for publishing data papers

- <https://www.nature.com/sdata/>
- <https://www.journals.elsevier.com/data-in-brief>
- <https://www.mdpi.com/journal/data>
- <https://academic.oup.com/gigascience>
- <https://openhealthdata.metajnl.com/>
- <https://bmccresnotes.biomedcentral.com/>
- <https://peerj.com> (not explicit for data papers, has to be wrapped as a research article)
- ...

# Guideline for Open Science data holdings

## FAIR Data Principles:

- Findable – Digital Object Identifier ([DOI!](#))
- Accessible
- Interoperable – how external services can interact with
- Reusable

The FAIR Guiding Principles for scientific data management and stewardship  
<https://www.nature.com/articles/sdata201618>



# Open data & Ethics

- Sharing everything (protocols, data, programming codes, etc.) is the most ethical approach possible, but there is at least one big little problem...
- The privacy of who the data is!
- Therefore, before publishing your data (or using someone else's data), make sure there are no ethical issues with that data.

# Recommendations for open data science

1. Provide or cite source code for tools in public repositories
2. Provide or cite pipelines in public repositories
3. Train scientists in data science
4. Journal editors and reviewers must enforce computational reproducibility

[https://en.wikipedia.org/wiki/Open\\_science\\_data](https://en.wikipedia.org/wiki/Open_science_data)

<https://gigascience.biomedcentral.com/articles/10.1186/s13742-016-0127-4>

# Example

**Open data in Biomechanics and Motor Control**

# Datasets in Biomechanics and Motor Control

Open data is not new in BMC, e.g.:

- Winter's book 1st ed (1979)
- <http://www.clinicalgaitanalysis.com/data/> (~2000)


A curated list of large datasets in BMC:

- <https://github.com/modenaxe/awesome-biomechanics#datasets-dvd>

# Open Science Data & BMClab

BMClab datasets – BMClab

← → ↻ ↗ ⚠ Not secure pesquisa.ufabc.edu.br/bmclab/datasets/ ☆ 🔒 📄 ⚙ 📁 👤 ⋮



# BMClab


## Biomechanics and Motor Control Laboratory

BMClab People Teaching Research Services Resources Publications BLOG

### BMClab datasets

The **BMClab** is committed to *open science data*.


[Click here to go to figshare to have access to our datasets or choose below a dataset.](#)



#### BDS: A public data set of human balance evaluations


This data set comprises signals from the force platform (raw data for the force, moments of forces, and centers of pressure) of 163 subjects plus one file with information about the subjects and balance conditions and the results of the other evaluations. Available at [PhysioNet](#) (DOI: 10.13026/C2WW2W) and at [figshare](#) (DOI: 10.6084/m9.figshare.3394432).


View page »



#### RBDS: A public data set of running biomechanics

Fukuchi RK, Fukuchi CA, Duarte M. (2017) A public dataset of running biomechanics and the effects of running speed on lower extremity kinematics and

 Português

 English

Search

Search

• [BMClab sitemap](#)

• [Useful links](#)

Events @ UFABC

• [BMClab agenda](#)

• [Seminars in Biomedical Engineering](#)

• [Seminars in Neuroscience](#)

• [All events](#)

# BMClab datasets @ figshare

The screenshot shows a web browser displaying the 'BMClab datasets' collection on Figshare. The URL in the address bar is [https://figshare.com/collections/BMClab\\_data\\_sets/3257938/8](https://figshare.com/collections/BMClab_data_sets/3257938/8). The page indicates '6 results found' and is sorted by 'Date added'. The datasets are displayed in a grid of six items, each with a 'DATASET' icon, a title, a posting date, and the author's name with a dropdown arrow.

Dataset Title	Posted On	Author
A public dataset on long-distance running training in 2019 and...	15.09.2021	Leonardo Afonseca
Dataset of gait and inertial sensors	17.03.2020	Desiree Miraldo
A public data set of overground and treadmill walking...	16.07.2018	Claudiane Fukuchi
A public data set of running biomechanics and the effects of...	22.03.2017	Reginaldo Fukuchi
A data set with kinematic and ground reaction forces of...	27.02.2017	Damiana A. dos Santos
A public data set of human balance evaluations	24.06.2016	Damiana A. dos Santos

At the bottom of the page, there is a navigation bar with links: About, Features, Tools, Blog, Ambassadors, Contact, FAQs, Privacy Policy, Cookie Policy, T&Cs, Sitemap. Social media icons for Facebook, Twitter, and YouTube are also present. The footer includes the text 'figshare. credit for all your research.' and logos for 'Part of DIGITALSCIENCE', 'CHRONOPOLIS', 'DataCite', 'C O P S', and 'OPEN ACCESS'.

Duarte, Marcos; Fukuchi, Reginaldo; Watanabe, Renato Naville; Fukuchi, Claudiane; Miraldo, Desiree (2018): BMClab datasets. figshare. Collection.

<https://doi.org/10.6084/m9.figshare.c.3257938.v8>

# In conclusion ...

- I hope what I have said today is useless for you ...
- ... this time because the next generation of thinkers (you) consider absurd that on the quest for the well-being of society, we hide information from each other.
- Obrigado.