# 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 ...

#### 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: <a href="https://www.nature.com/sdata/policies/repositories">https://www.nature.com/sdata/policies/repositories</a>

## 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://bmcresnotes.biomedcentral.com/
- <a href="https://peerj.com">https://peerj.com</a> (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

#### **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

Example

Open data in Biomechanics and Motor Control

#### **Datasets in Biomechanics and Motor Control**

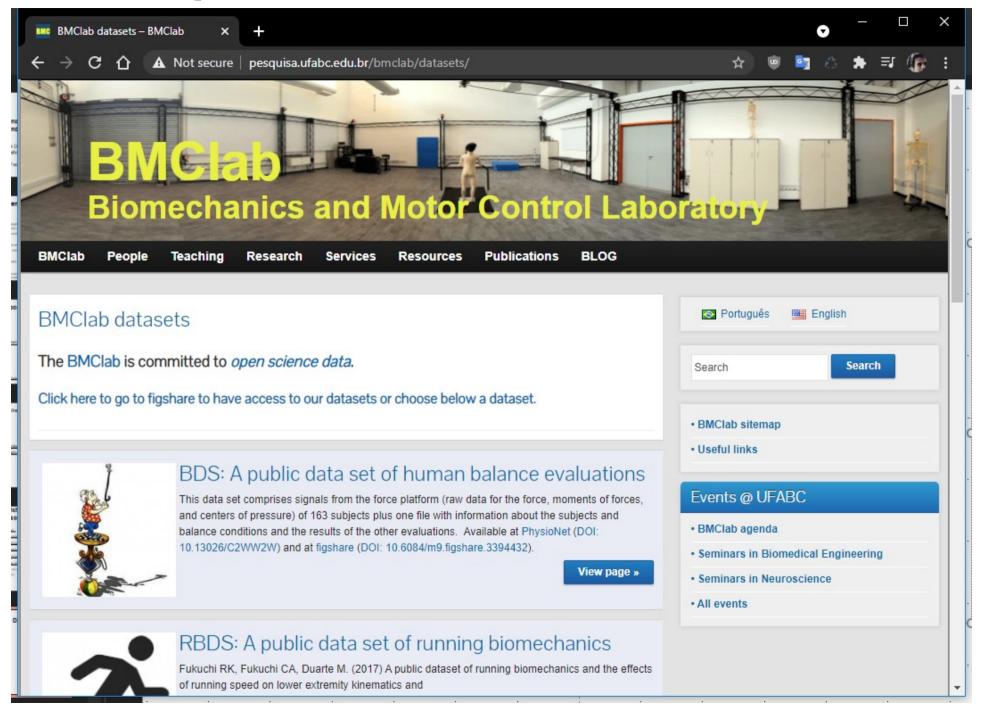
Open data is not new in BMC, e.g.:

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

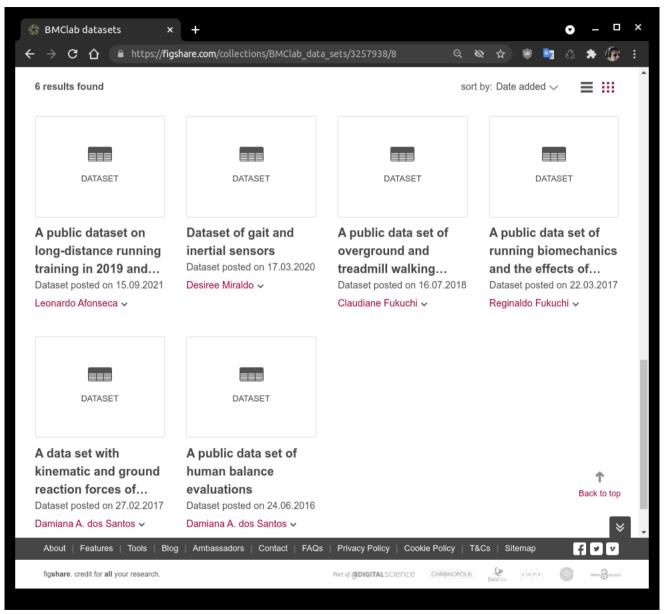
A curated list of large datasets in BMC:

 https://github.com/modenaxe/awesome-biomechanics#dat asets-dvd

## **Open Science Data & BMClab**



## BMClab datasets @ figshare



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 wellbeing of society, we hide information from each other.
- Obrigado.