R in Grenoble
DATA CHALLENGES

Magali Richard & Florent Chuffart

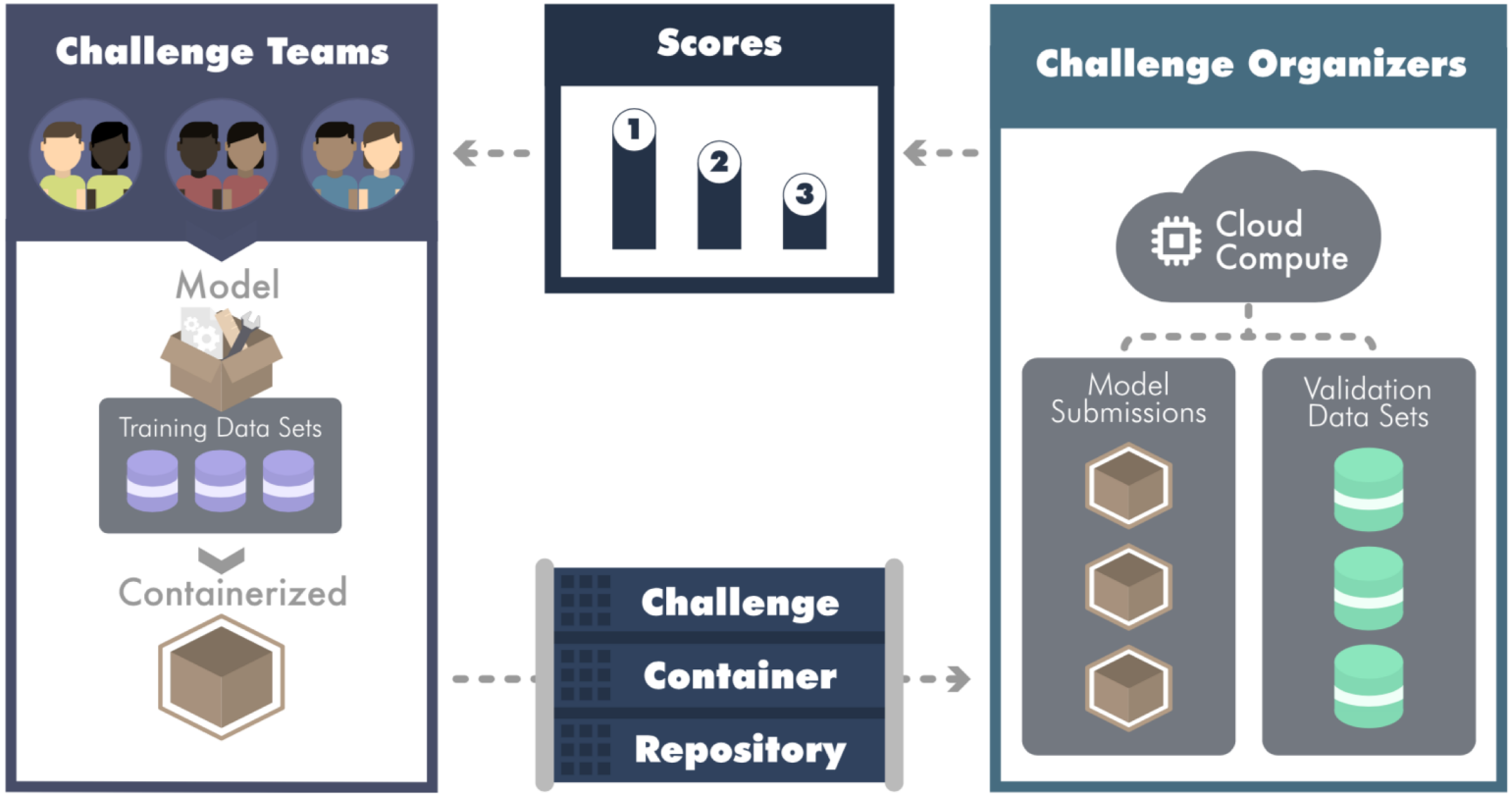
Introduction

Data challenges in class

Data challenges for scientists

Tutorial

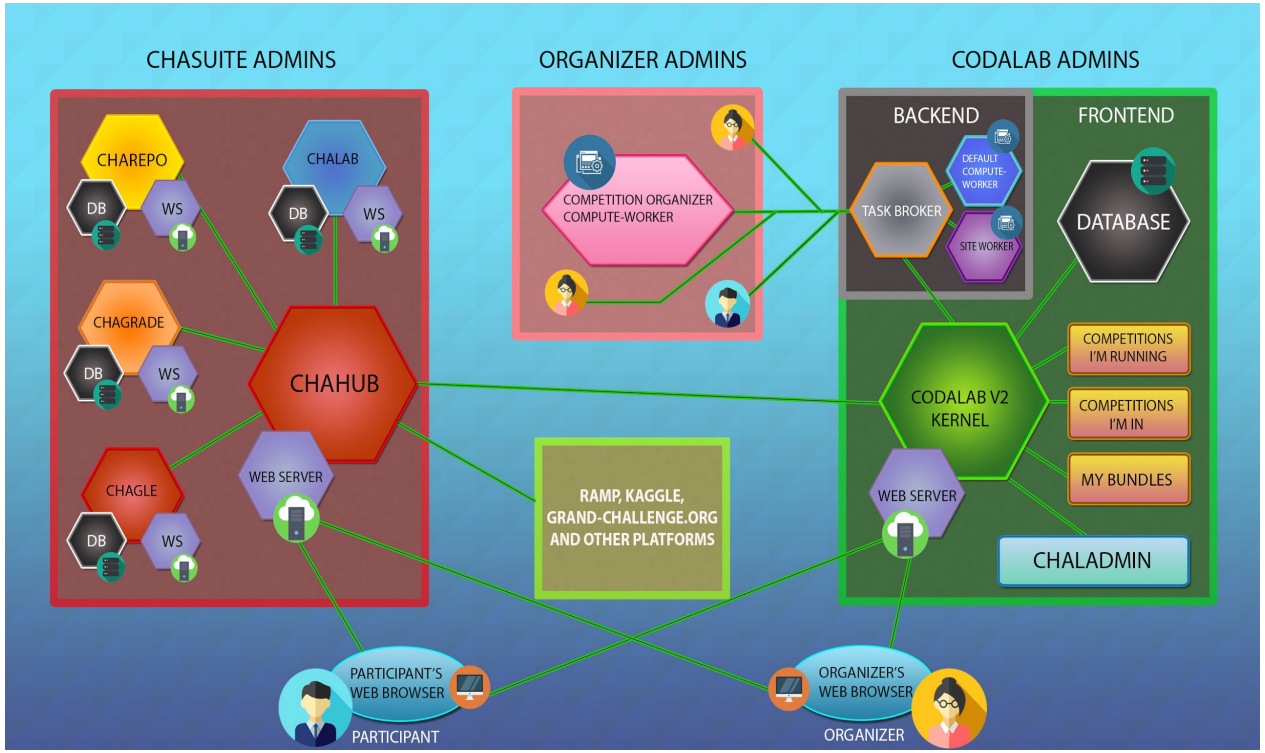
What is a data challenge?



Ellrott et al. Genome Biology (2019) 20:195

The challenge platform

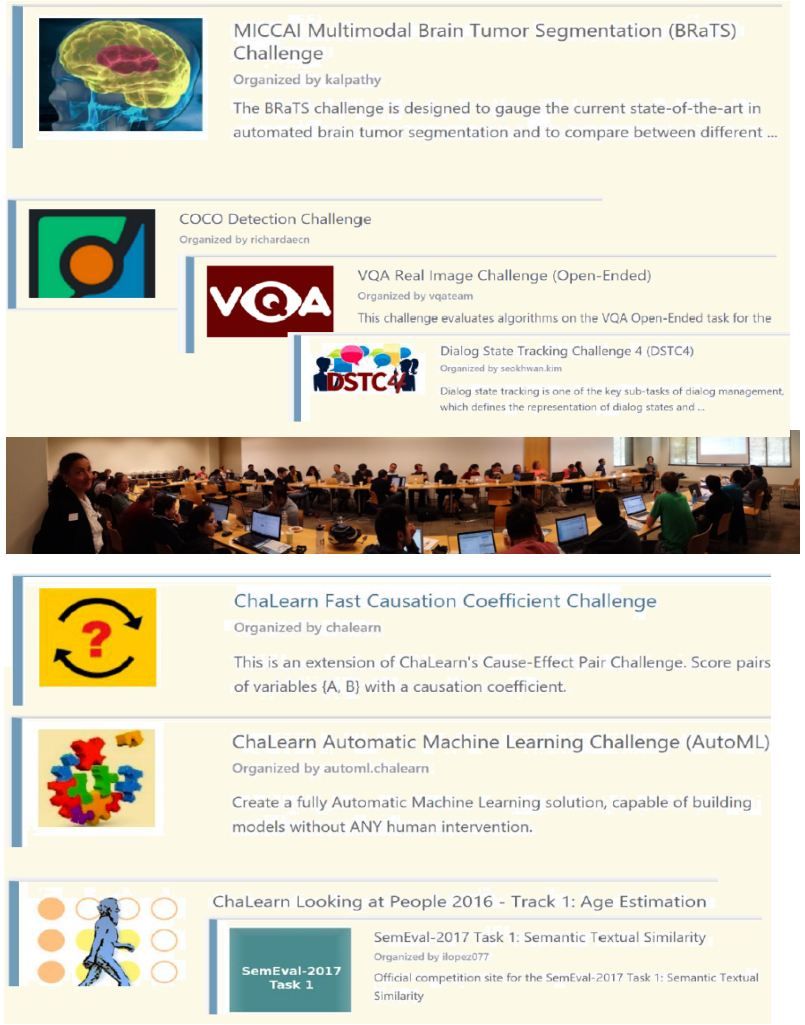
CodaLab



- Enables participants to submit their codes
- Automatically rank the participants

Alexis Arnaud

History



2013: Medical data. Result submission.
COMPETITION BUNDLES

2014: Computer vision, speech, NLP, IR.
MSCOCO: 361 participants.

2015: **CODE SUBMISSION**
AutoML: 687 participants
Hackathons. Coopetitions.

2016: ChaLab Wizard.
USE IN EDUCATION

2017: 480 challenges, 10000 users.
SCALABILITY, REUSABILITY,
DYNAMIC COMPETITIONS
See.4c: EU prize, 2 million Euros

Introduction

Data challenges in class

Data challenges for scientists

Tutorial

A novel pedagogic approache

As practical work

Sex Prediction Challenge (sexpred)

Organized by Alexis_Arnaud

The challenge provides the data.frame d of tumoral tissues described by genes expression values, histological and clinical attributes. The ...

Nov 15, 2019-Dec 30, 2019

37 participants

As homework

Histology Prediction Challenge (histpred)

Organized by Alexis_Arnaud

The challenge provides the data.frame d of tumoral tissues described by genes expression values, histological and clinical attributes. The ...

Nov 21, 2019-Dec 30, 2019

37 participants

As final evaluation

Virulence Prediction Challenge (virpred)

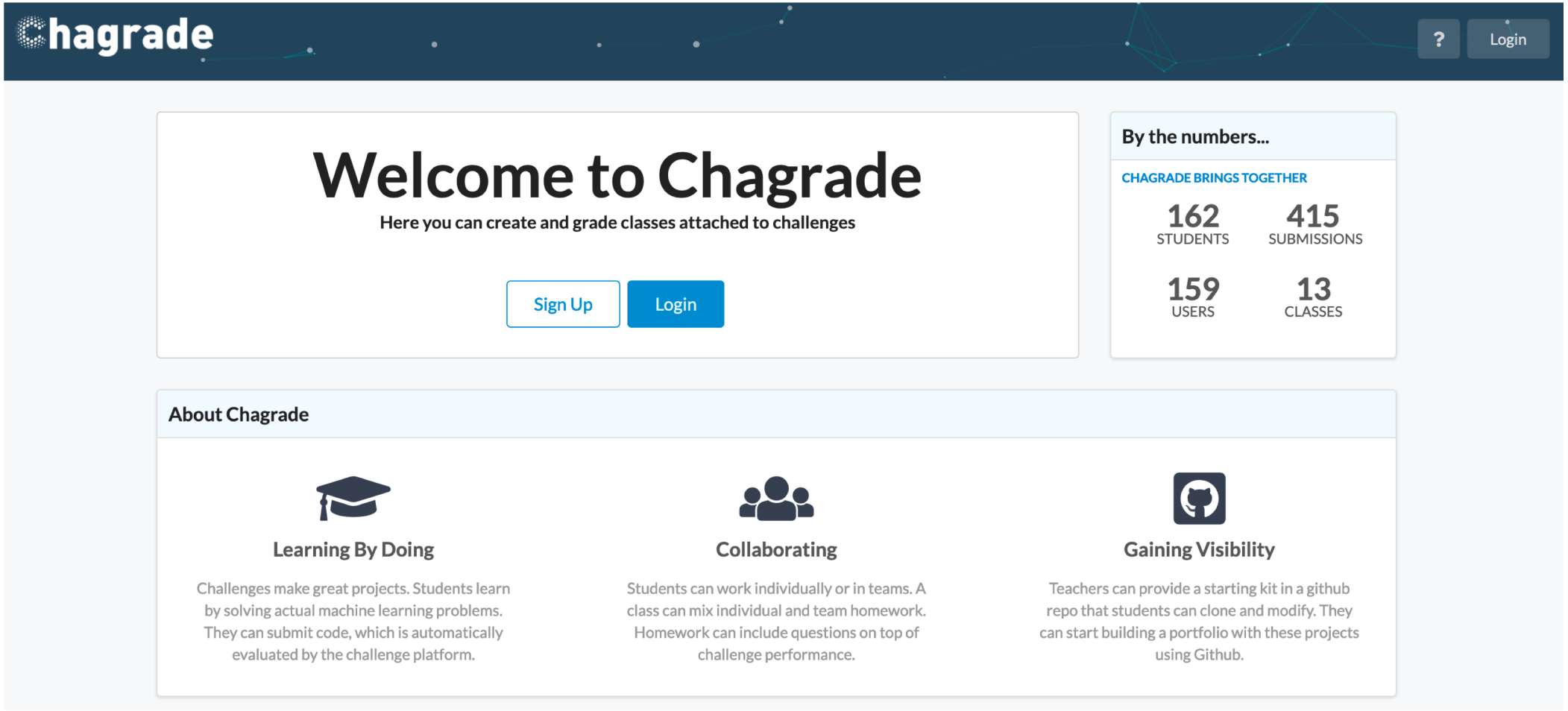
Organized by FlorentC

The challenge provides the data.frame d of tumoral tissues described by genes expression values, histological and clinical attributes. The ...

Jan 09, 2020-Jan 11, 2020

35 participants

Chagrade: a dedicated tool in codalab



Introduction

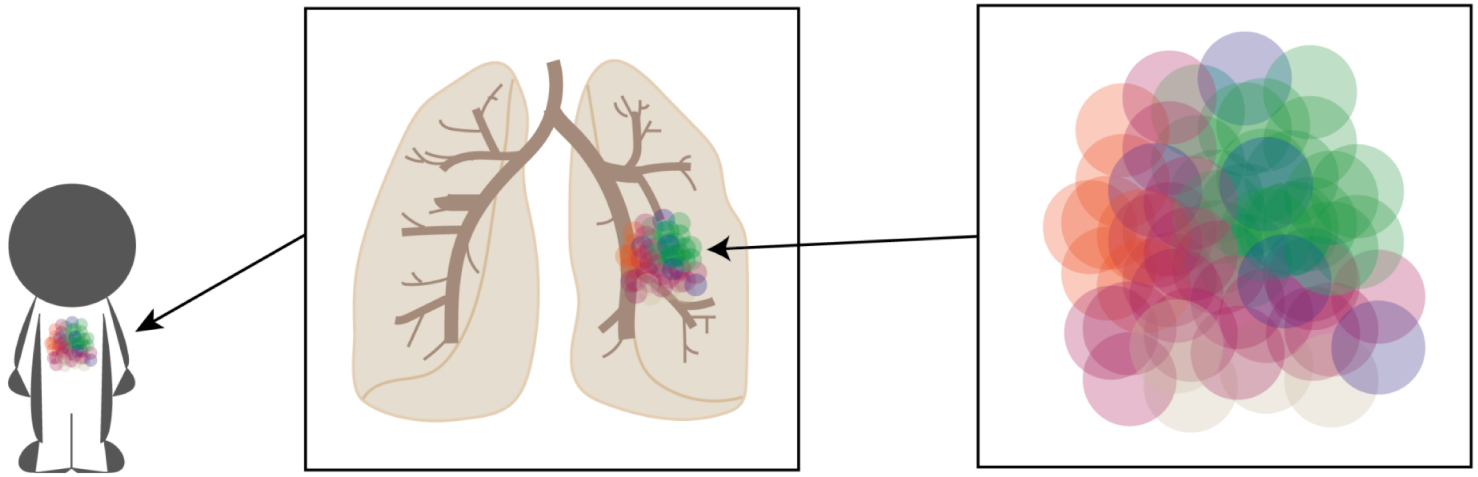
Data challenges in class

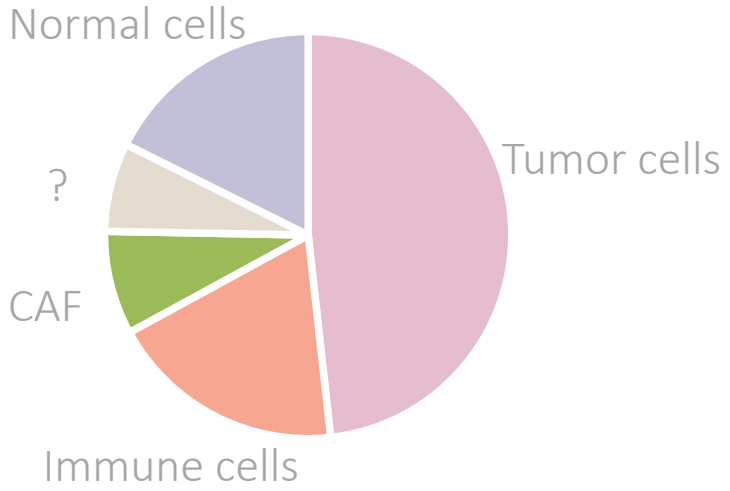
Data challenges for scientists

Tutorial

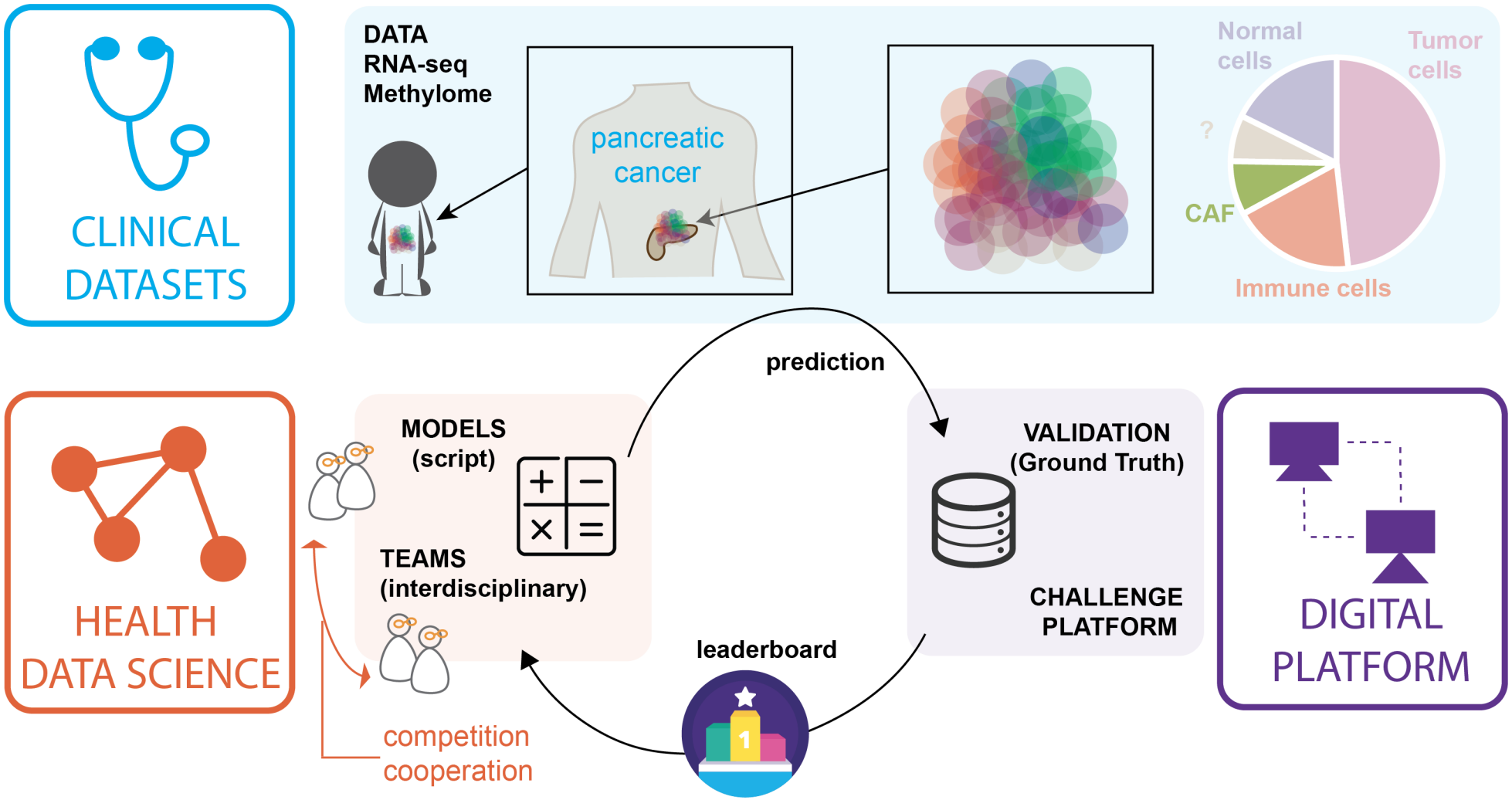
Goal of our data challenge

- **Quantification of tumor heterogeneity**



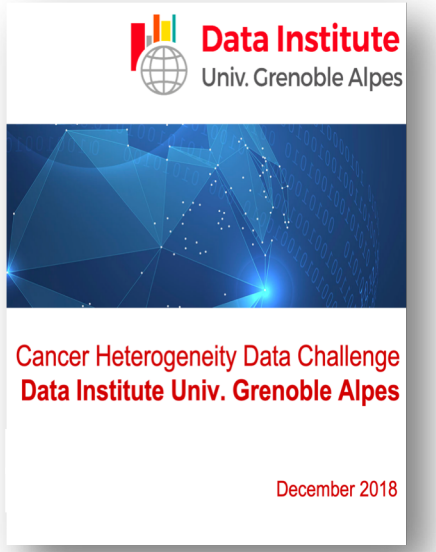


A challenge for scientists



2 editions of the data challenge

1rst edition (2018)



UNIVERSITÉ Grenoble Alpes

eit Health

- Methylation Data
- One cancer type
- Cell lines

2nd edition (2019)

UNIVERSITÉ Grenoble Alpes

eit Health

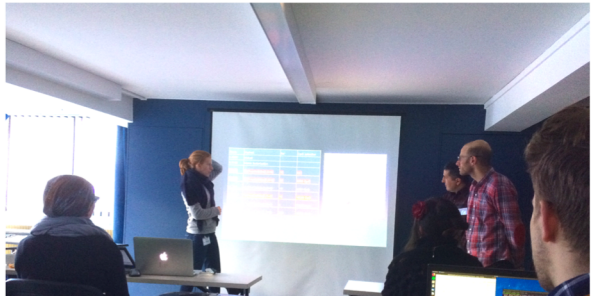
CIT CARTES D'IDENTITÉ DES TUMEURS

LA LIGUE CONTRE LE CANCER

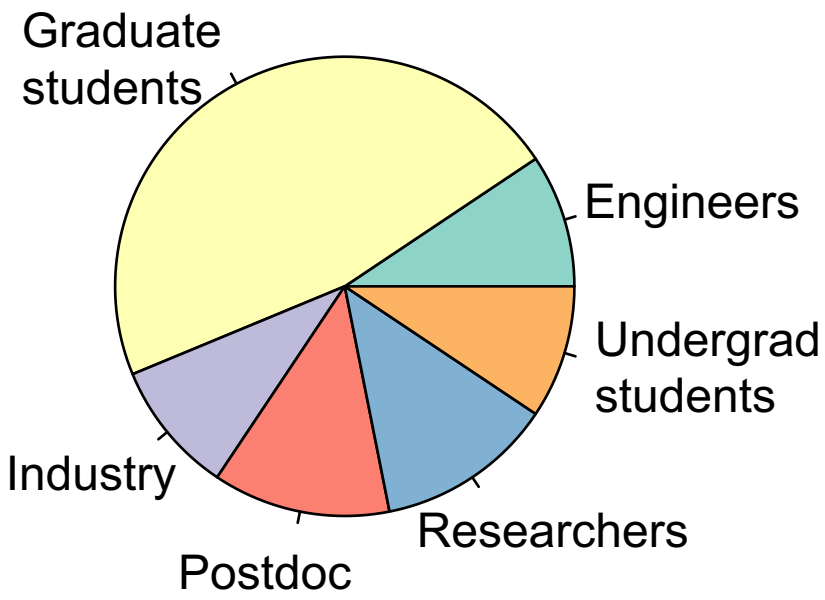
- Methylation and transcriptomic Data
- Several cancer types
- Primary tumors / cell lines

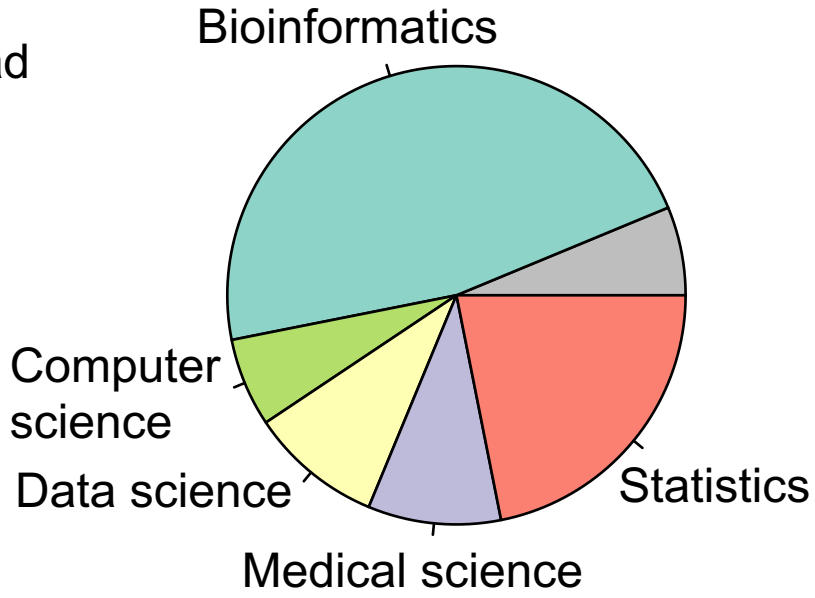
When and where?

- **When**: 25-29 of November 2019
- **Where**: Aussois (CAES CNRS), French Alps

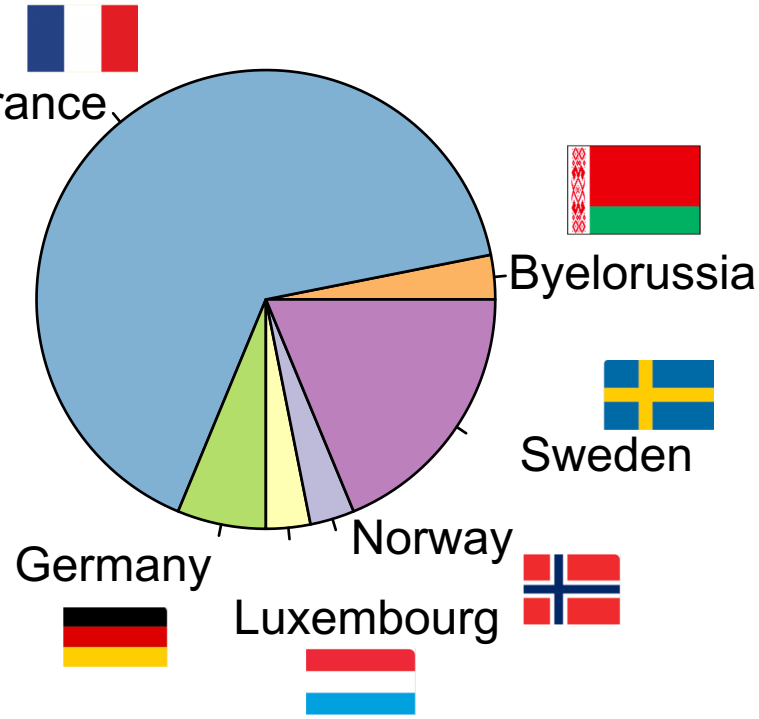


About the participants (n=34)



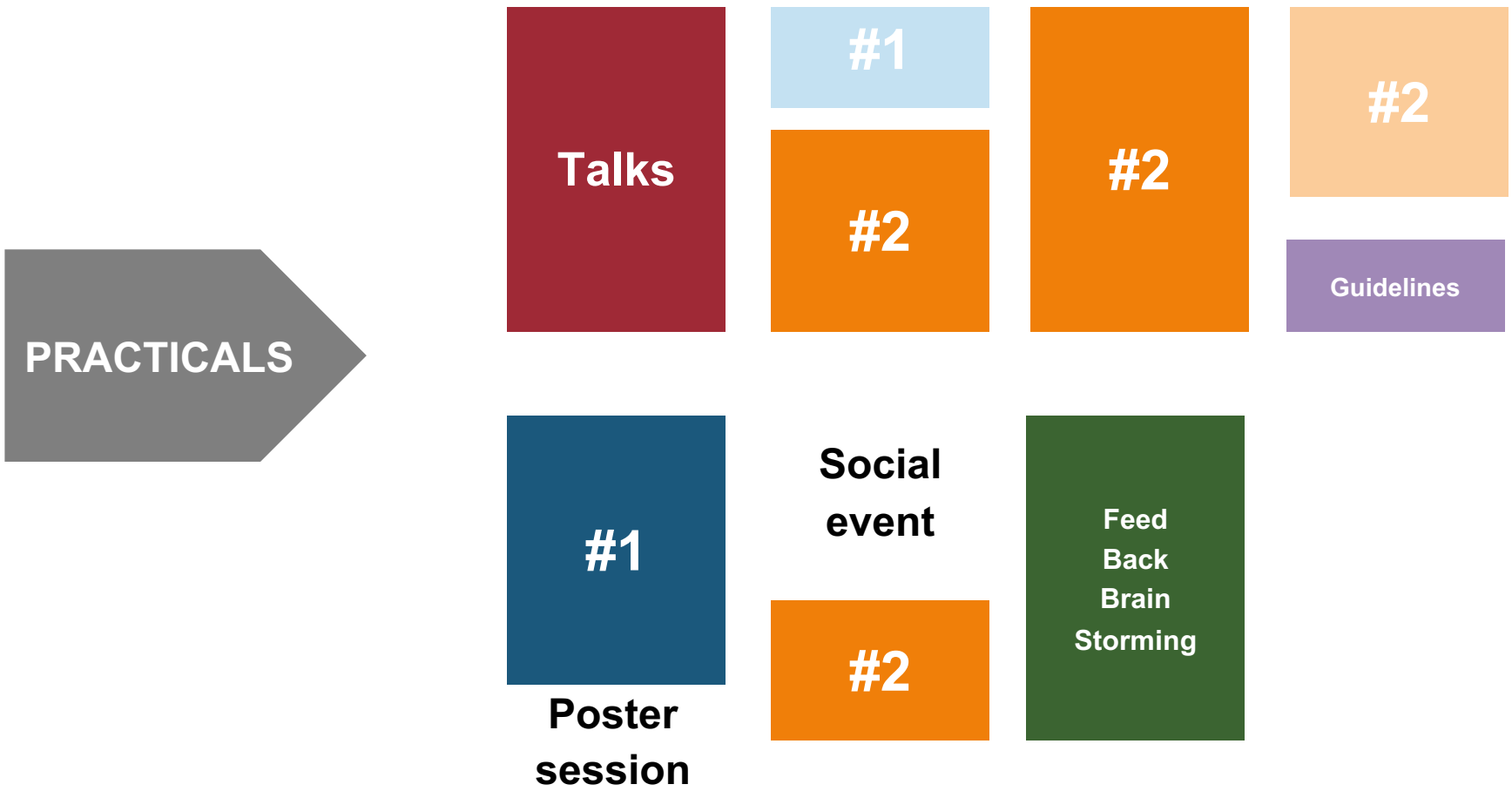


Institut Curie
UGA Grenoble
CRC Cordeliers
INSERM
CEA
Innate Pharma
Verteego
…



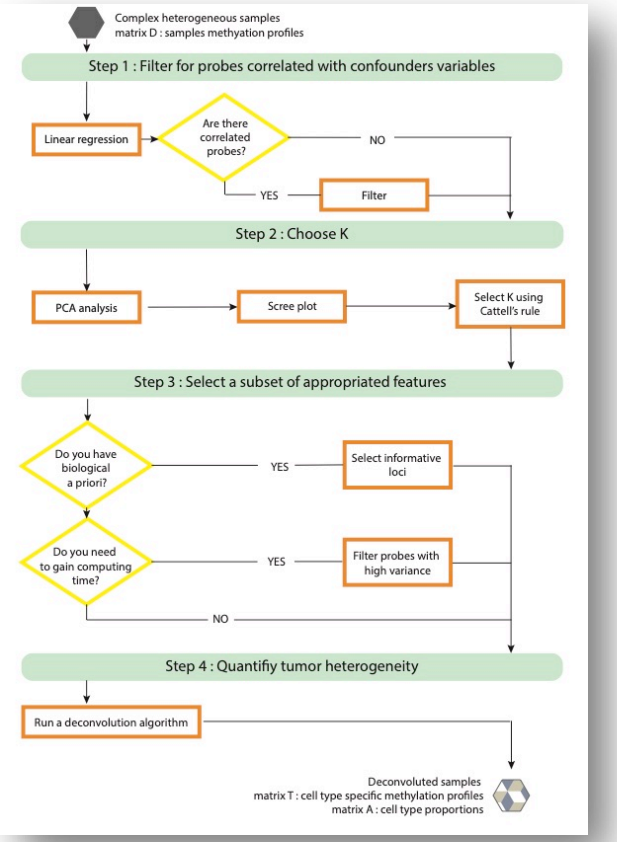
10 teams of 3-4 people

Agenda

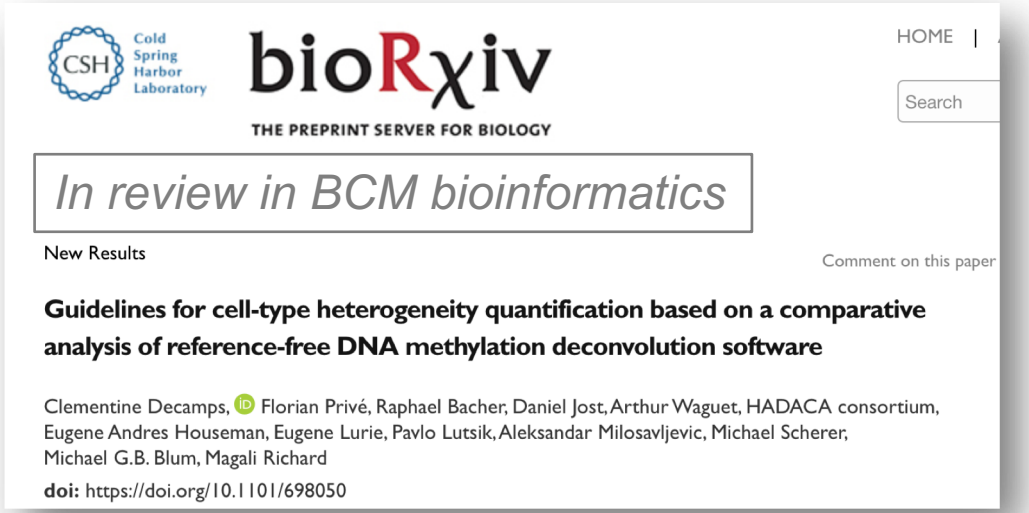

PRACTICALS
Talks
#1
#2
#2
#2
Guidelines
#1
Poster session
Social event
#2
Feed Back Brain Storming

Valorization of the first edition

➢ Guidelines



➢ Article



➢ R package *medepir*

https://rdrr.io/github/bcm-uga/medepir/man/medepir-package.html

M Richard, C Decamps , F Privé, M Blum

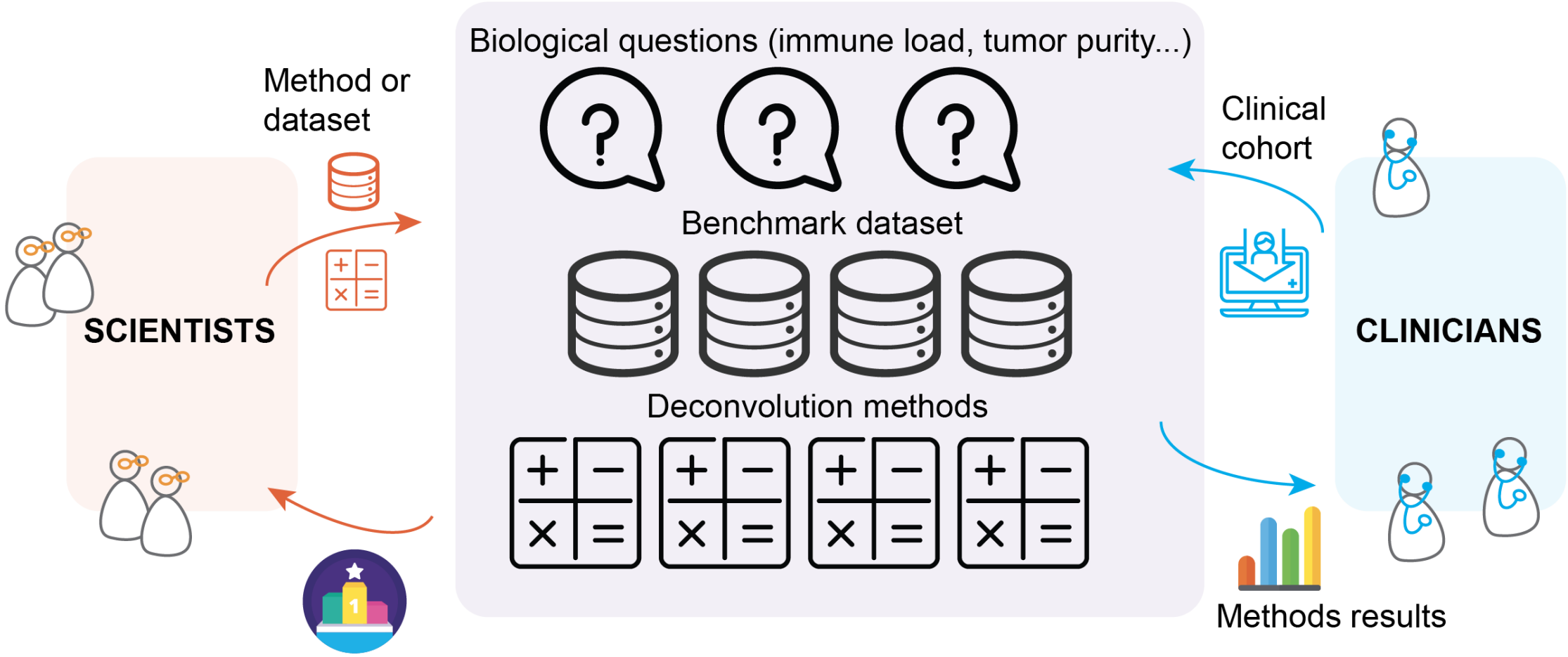
➢ Blog posts

Health data challenges organization: feedback, comments and recommendations.

Authors: Elise Amblard, Yuna Blum, Jane Merlevede, Magali Richard

In preparation

From data challenges to benchmarks



Introduction

Data challenges in class

Data challenges for scientists

Tutorial

On line tutorial by Alexis Arnaud

Codalab › Bundle_generation › **Details**

Bundle_generation
Project ID: 5596

Star 0 | Fork 0 | Clone

11 Commits | **1** Branch | **0** Tags | **748 KB** Files

README.md

To create a Codalab competition, you have to generate a zip file with the required files inside and upload the archive on the Codalab platform. The *R* script `Bundle_generator_-_cmd.Rmd` allows you to automatically create a valid bundle by using the *R* terminal to ask you some questions in order to personalize the competition.

Generate a Codalab bundle

In a *R* terminal, run the following command : `rmarkdown::render(input = "Bundle_generator_-_cmd.Rmd", envir = new.env( ) )`

You can also generate a toy bundle, with the parameters by default, with *Rscript* : `Rsript -e 'rmarkdown::render(input = "Bundle_generator_-_cmd.Rmd", envir = new.env( ) )'`

Test the Codalab bundle

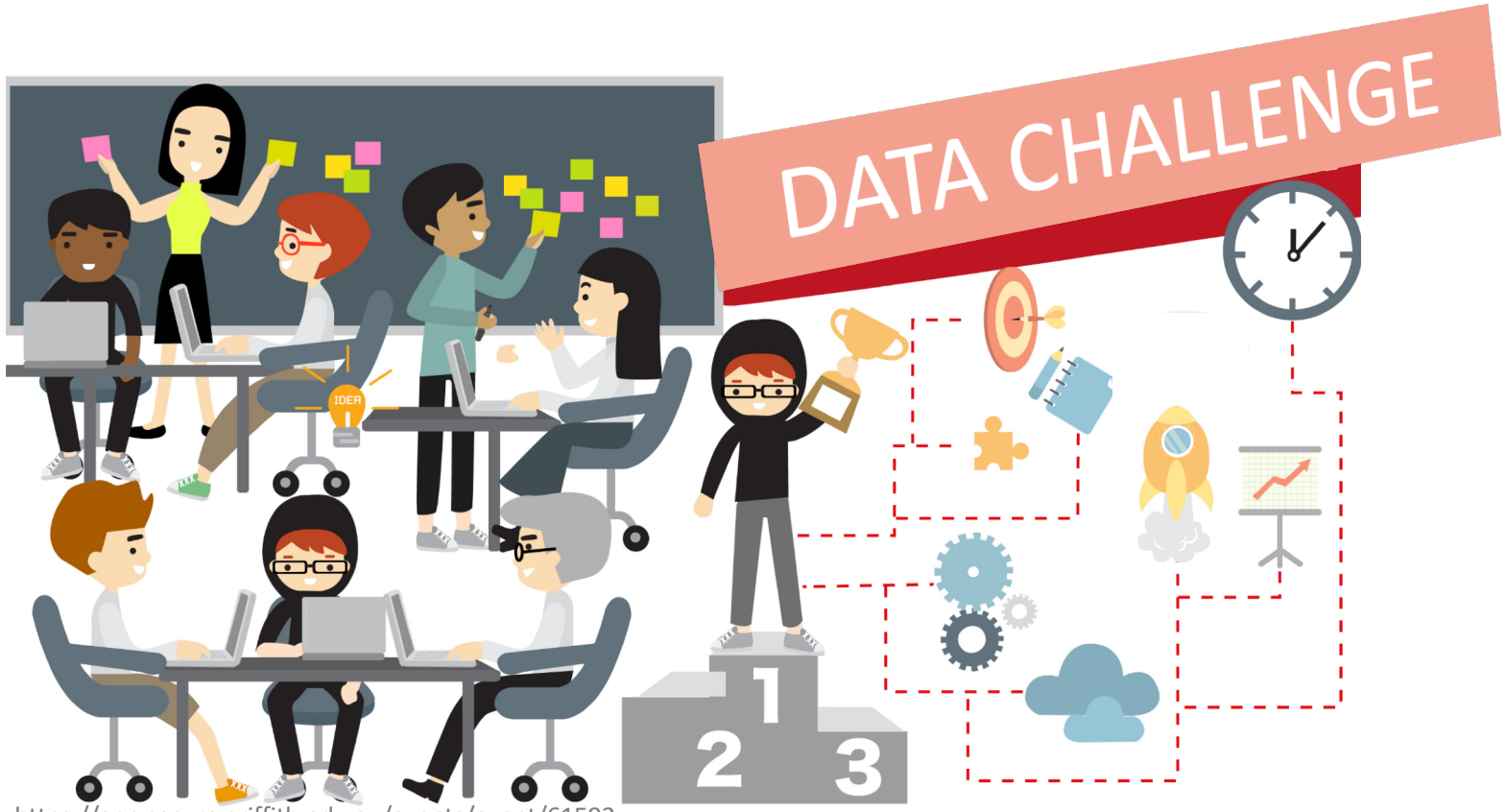
In a *R* terminal, run the following command : `rmarkdown::render(input = "Bundle_test.Rmd", envir = new.env( ) )`

[WIP] Render a web site to previsualize the competition

It gives you the possibility to visualize your competition with a minimalist website which looks like the Codalab platform.
`rmarkdown::render_site(input = "Web_page_generator.Rmd", envir = new.env( ) )`

It’s up to you now!

Thank you for your attention !



https://app.secure.griffith.edu.au/events/event/61593