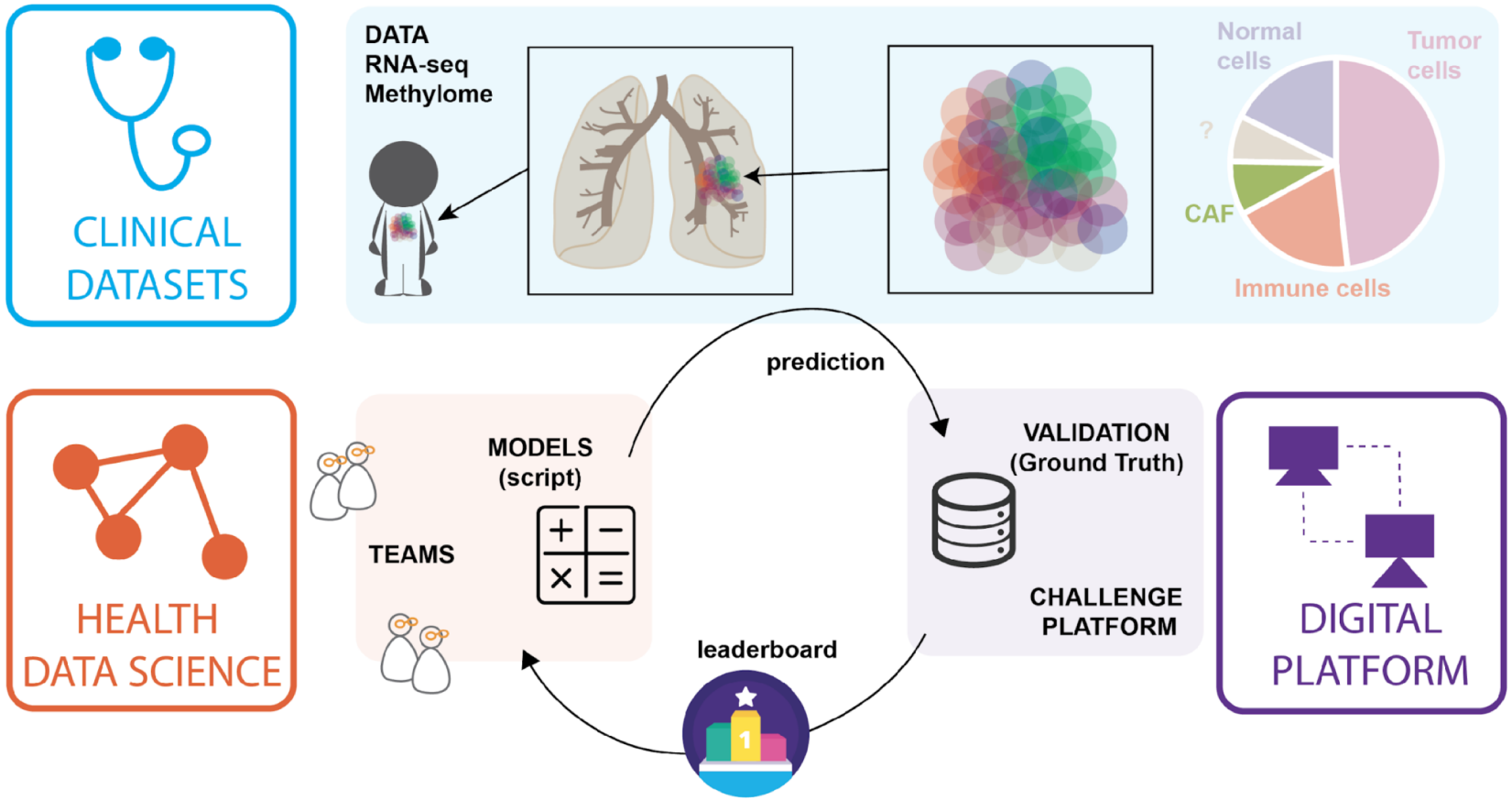
HADACA – Health Data Challenge

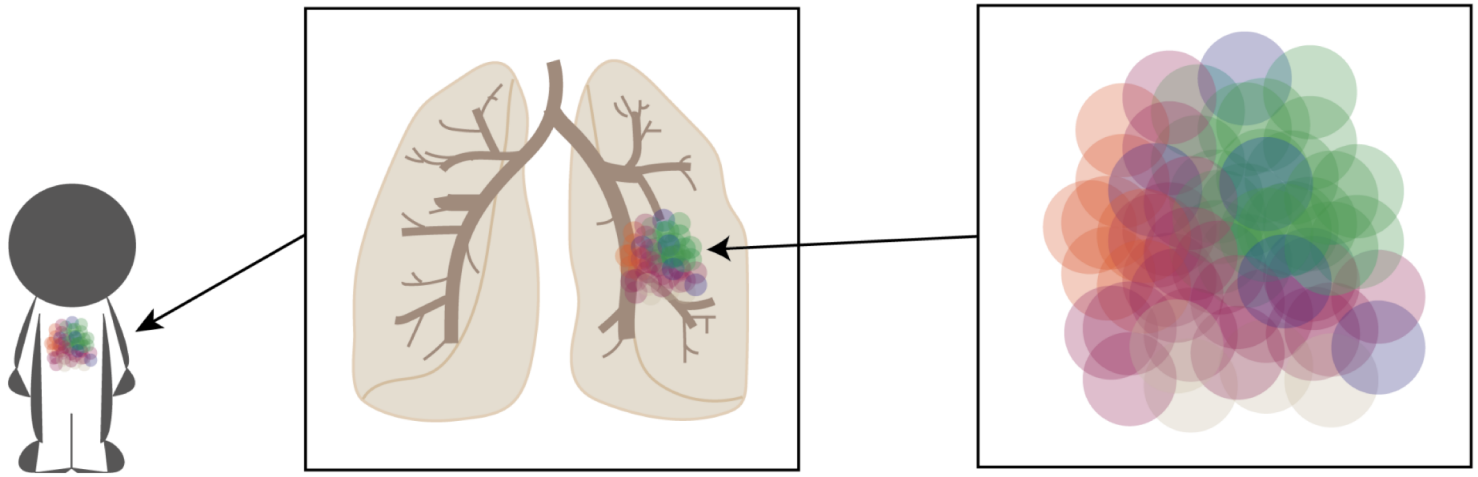
Deconvolution methods to quantify tumor heterogeneity

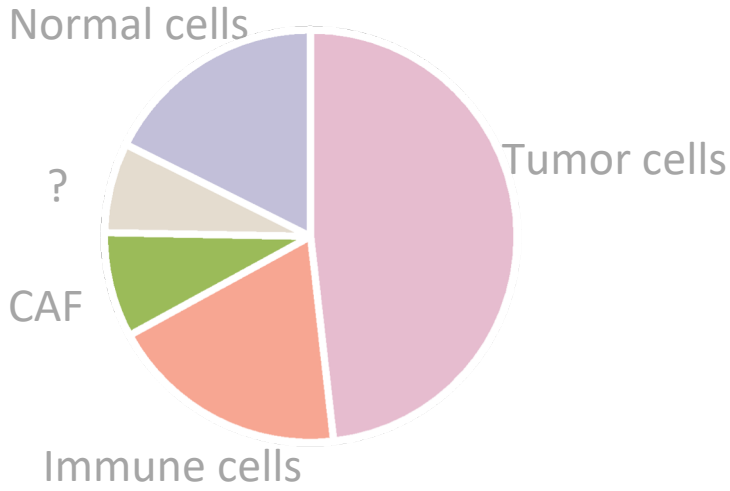
Introduction to the data challenge



Goal of our data challenge

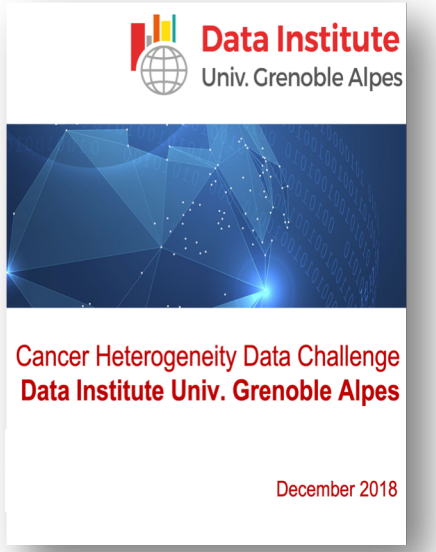
Quantification of tumor heterogeneity





2nd edition 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

Organizing team

Magali Richard
(TIMC-IMAG team)

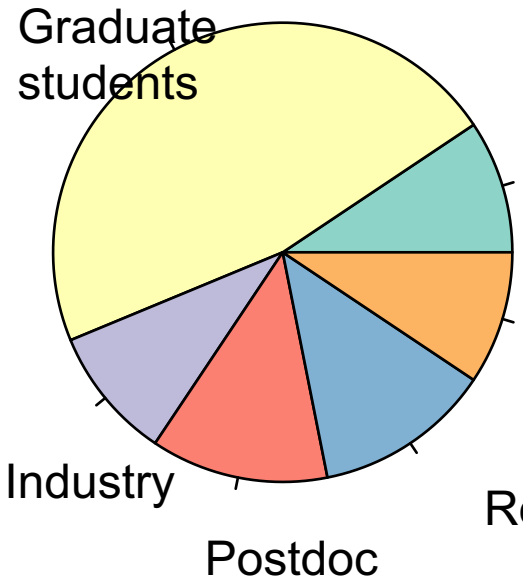
Clémentine Decamps
(TIMC-IMAG team)

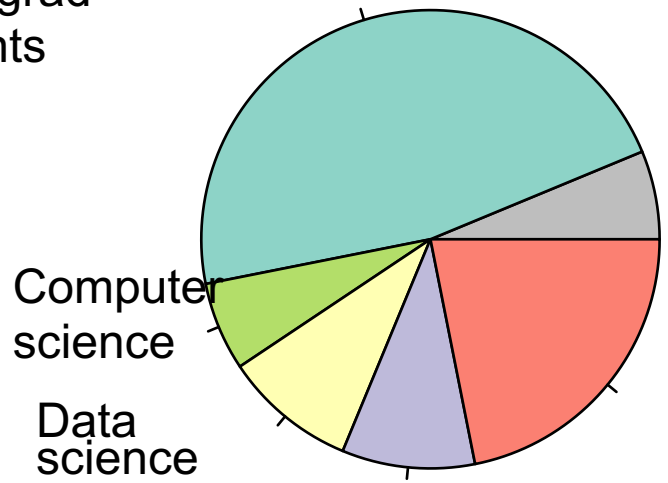
Alexis Arnaud
(TIMC-IMAG team)

Yuna Blum
(CIT team)

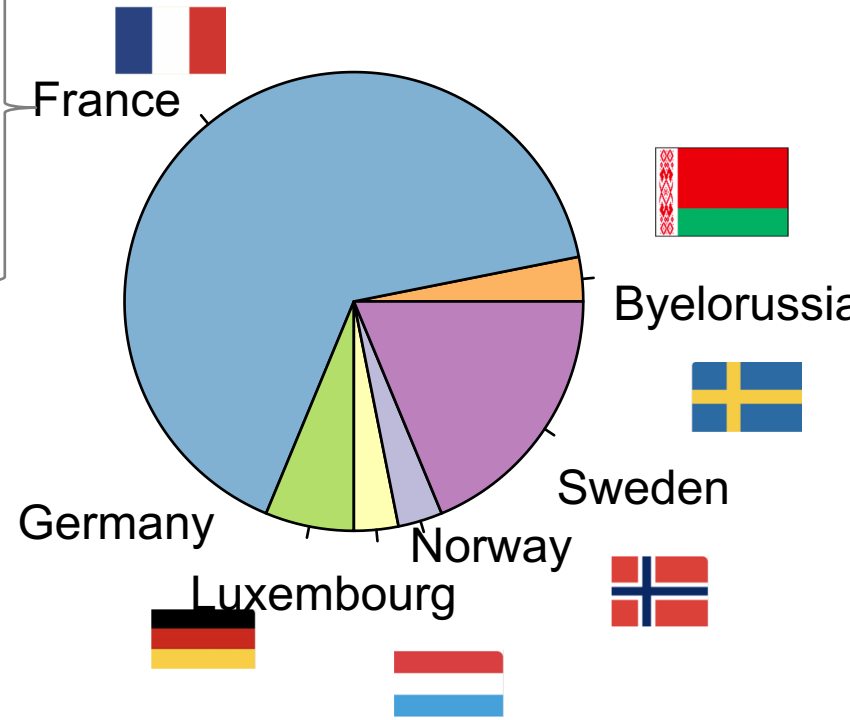
Florent Petitprez
(CIT team)

About the participants (n=31)





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



10 teams of 3-4 people

Invited speakers

Jérôme Cros
AP-HP, Paris
France

- Tumor heterogeneity: the clinician's point of view

Michael Scherer
Max-Planck-Institut für Informatik, Saarbrucken,
Germany

- Methylation and deconvolution methods

Francisco Avila Cobos
Ghent University, Gand
Belgium

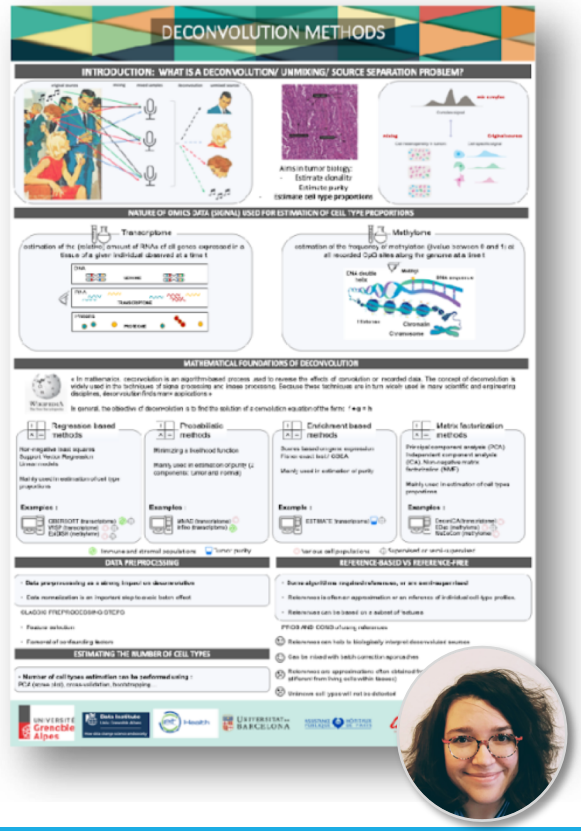
- Transcriptome and deconvolution methods

Agenda


DAY
1
Talks
Challenge #1
Poster session
DAY
2 #1
Challenge #2
Social event
#2
DAY
3
#2
Brain-storming
DAY
4
#2
Guidelines

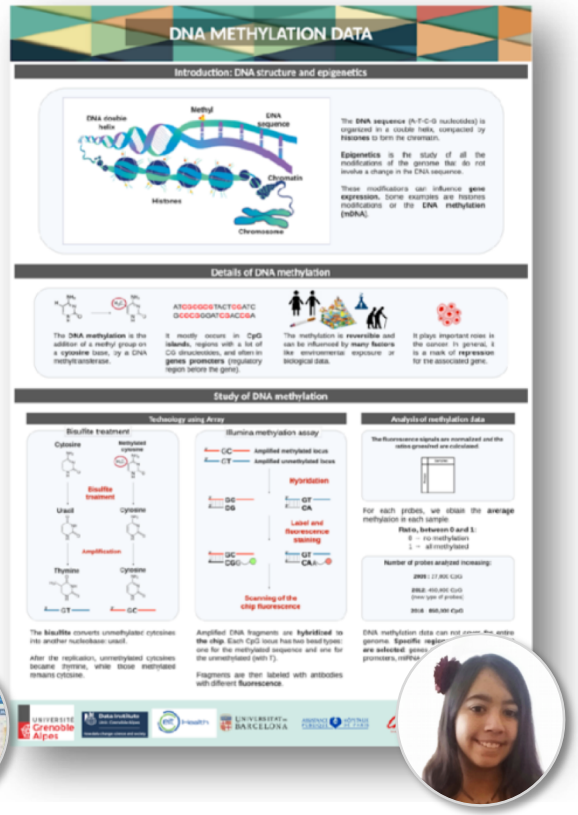
“The basics” during poster session

“Deconvolution methods”

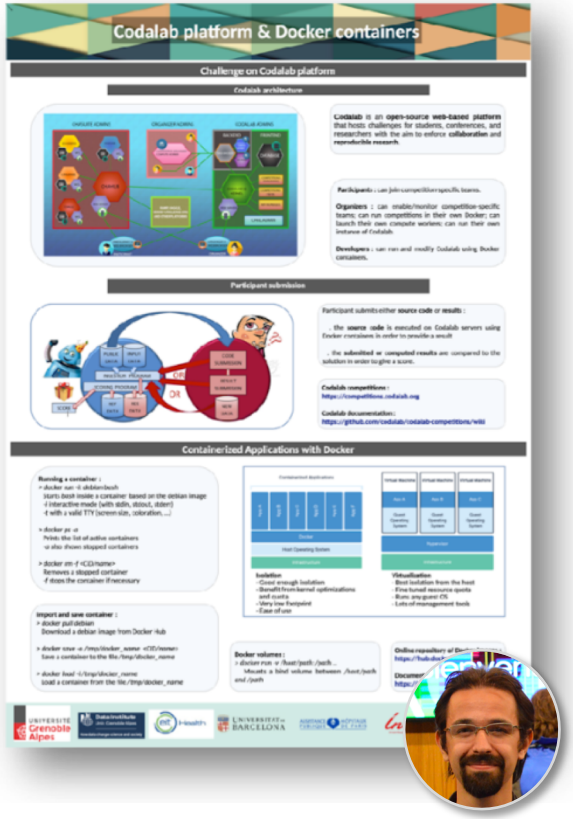


“Transcriptomic data”

“Methylome data”



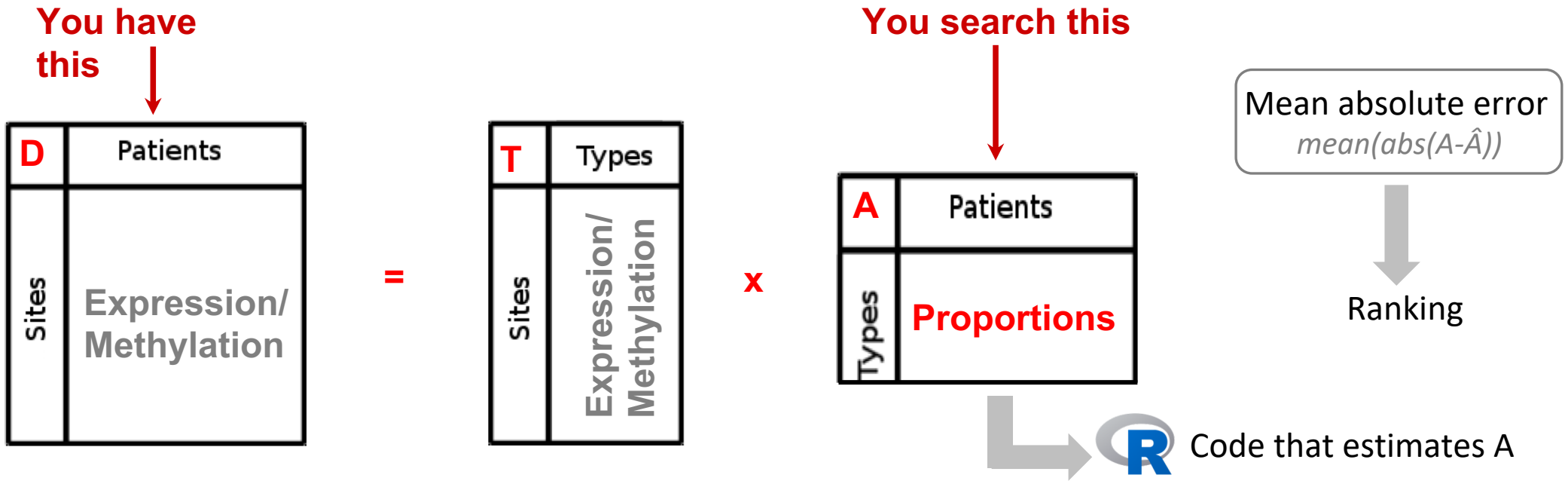
“Codalab platform & docker”



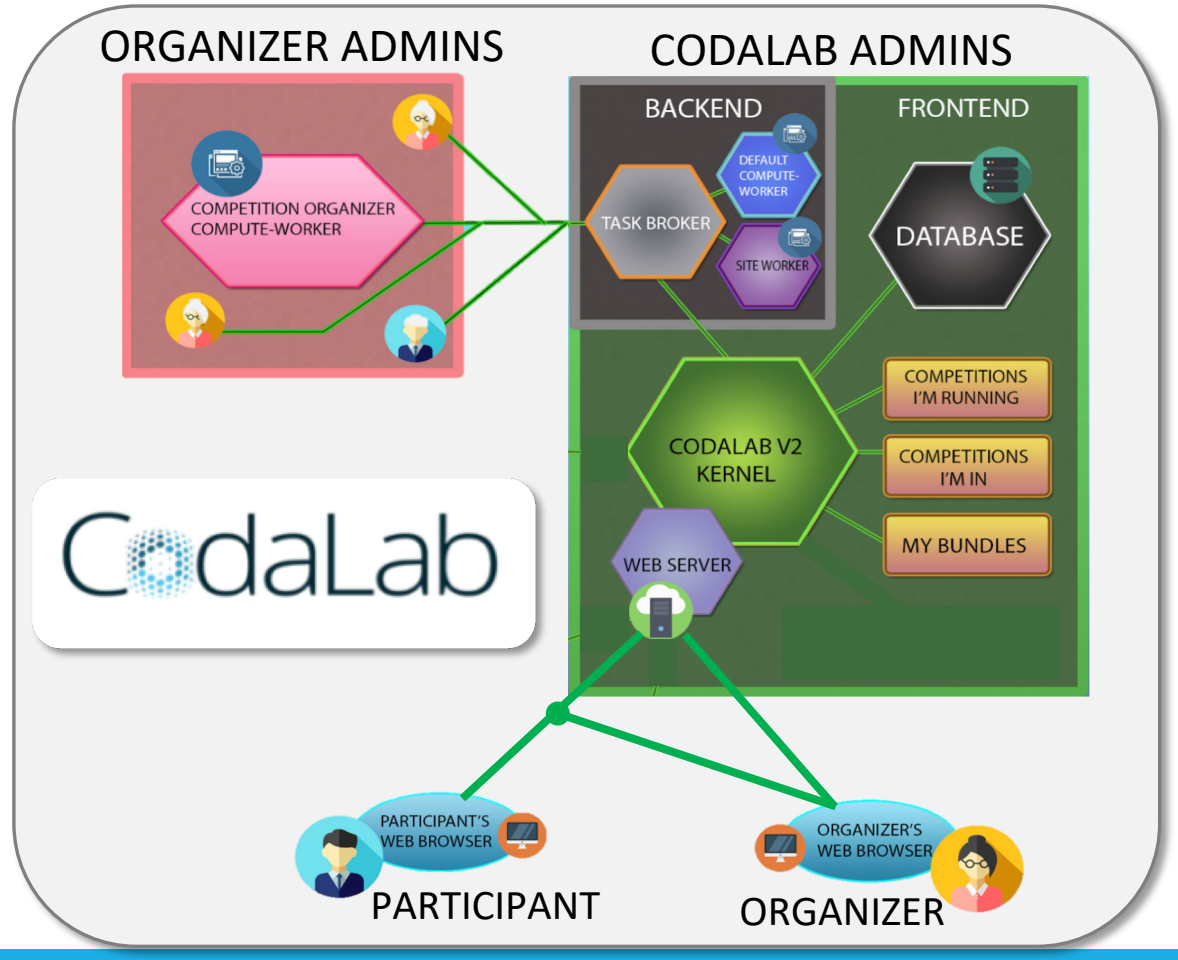
The challenge

Your goal: estimate the matrix A

The matrix A represents the proportion of each cell type in each patient



The challenge platform



Open-source platform

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

Alexis (IT) dedicated to the computing resource management

Challenges of increasing complexity

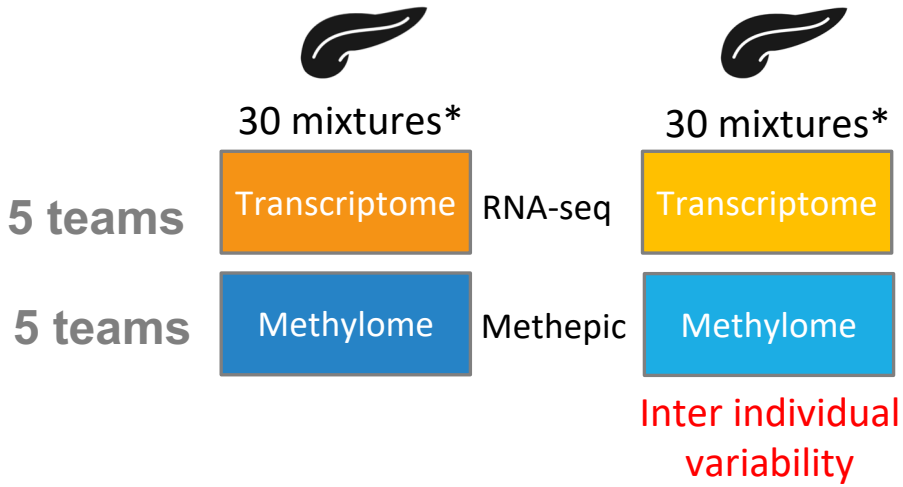
Challenge #1

Aim: learn how to use collalab, discover the dataset, manage to submit a deconvolution script on either RNA-seq or methylome

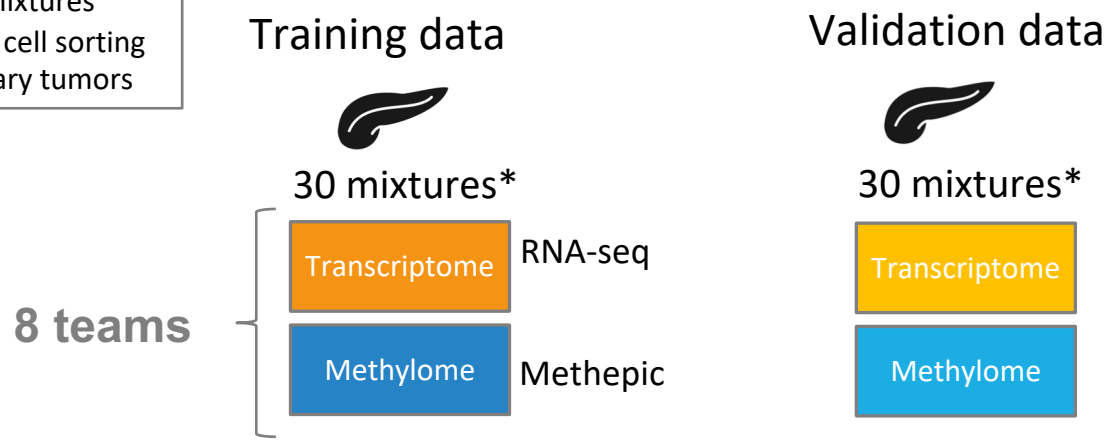
Challenge #2

More populations (some similar)

Aim: test and develop method to quantifiy tumor heterogeneity using both RNA-seq and methylome data



*In silico mixtures from FACS cell sorting from primary tumors



The awards

Winner of the main Challenge (#2)

Winner of the training Challenge (#1)

Best poster award

How to participate?

(1) **Register** to the challenge on Codalab
(2) **Find** your teammates
(3) **Download** the starting kit and the public dataset

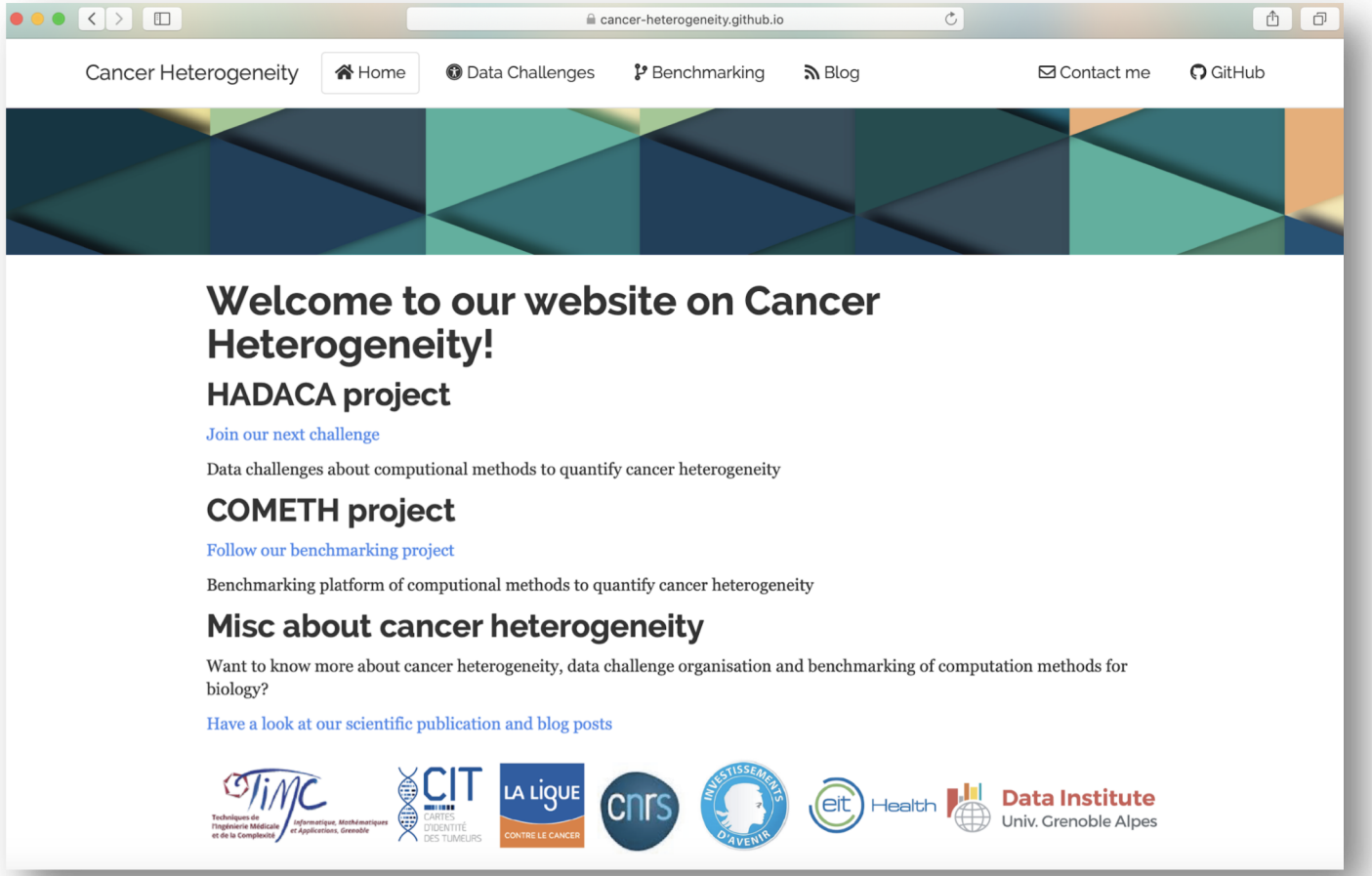
CHALLENGE BEGINS

(1) **Work** in group to find deconvolution methods
(2) **Submit** your results on the Codalab platform
(3) **Improve** your score

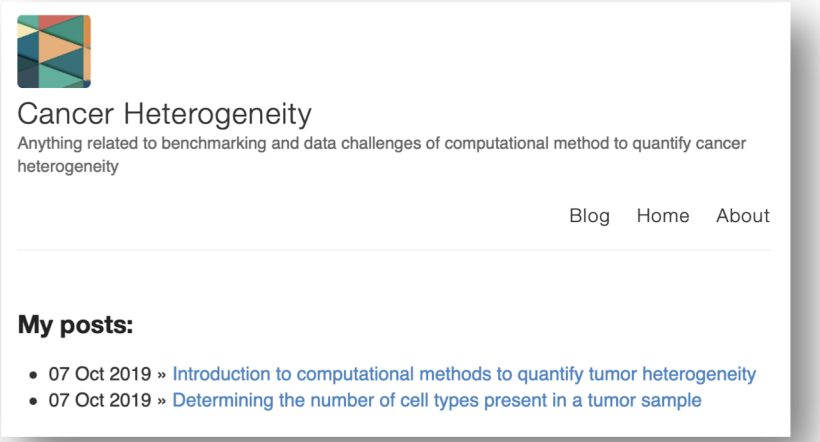
CHALLENGE ENDS

Restitution of your work (3 slides per team // PDF format)
at the end of each challenge (send it on time to Alexis!)

Website



- General information
- FAQ
- Blog posts
- List of methods
- …



Practical organization

- **Breakfast** : 7.30-9.15am (Level 4)
- **Dinner** starts at 7.45pm
- **Bar** open after lunch (Level 3) and after Dinner
  Coffee/tea offered (specify that you are from the Data Challenge)
- **Breaks** will take place in the bar
- **The poster session** will be in the mezzanine (Level 6)
  Poster numbers are in the program
- **For working sessions**, use the mezzanine or the 'La Scolette' room
- **Hike** will be on Wednesday afternoon (2 groups)

Objectives of the week

Share interdisciplinary knowledge

Learn good coding practices

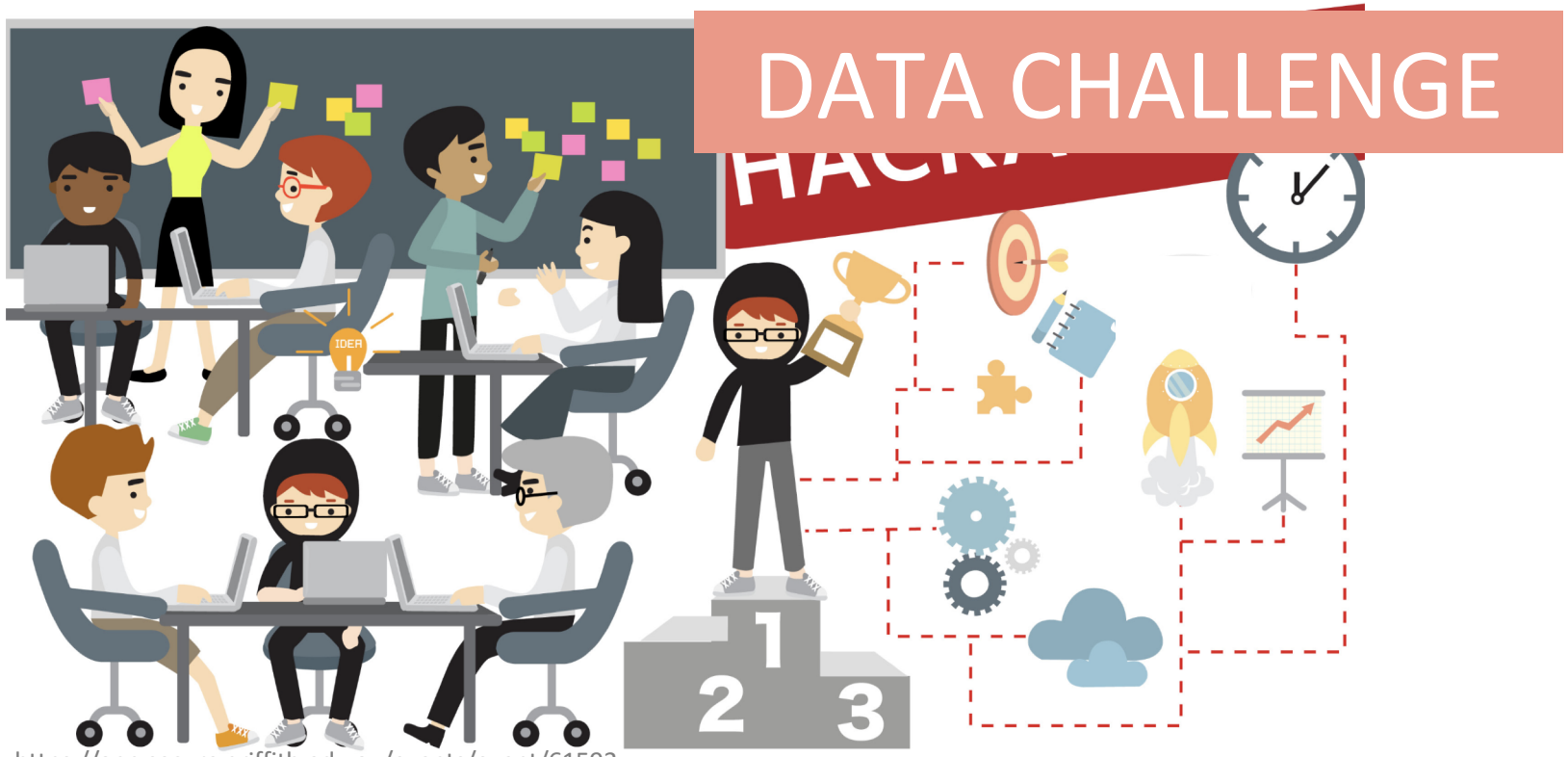
Discover methylome and transcriptomes specificities

Assess state of the art of deconvolution methods

Have fun

See you tomorrow at 9am, salle ‘La Scolette’, (or tonight at the bar)

Thank you for your attention !



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