



Karolinska
Institutet

SciLifeLab



Knowledge through bioinformatics

10 / 09 – 2019

Matthias Stahl

matthias.stahl@ki.se

Hej och välkomna!

Matthias heter jag
och
jag jobbar som *forskare*
på
Karolinska Institutet.



Bioinformatics

Data visualization

School

Informatics

Biochemistry & proteomics



A microscopic image of a blood smear stained with Wright-Giemsa. The field is dominated by numerous large, immature white blood cells, likely blasts, which have large, round to oval nuclei with dense, dark purple chromatin and scant, light blue cytoplasm. These cells are distributed across the field, with some appearing in small clusters. A few smaller, more mature-appearing cells are also visible, particularly towards the right side of the image. The background is a pale pinkish-white, representing the plasma or serum.

Acute Leukemias

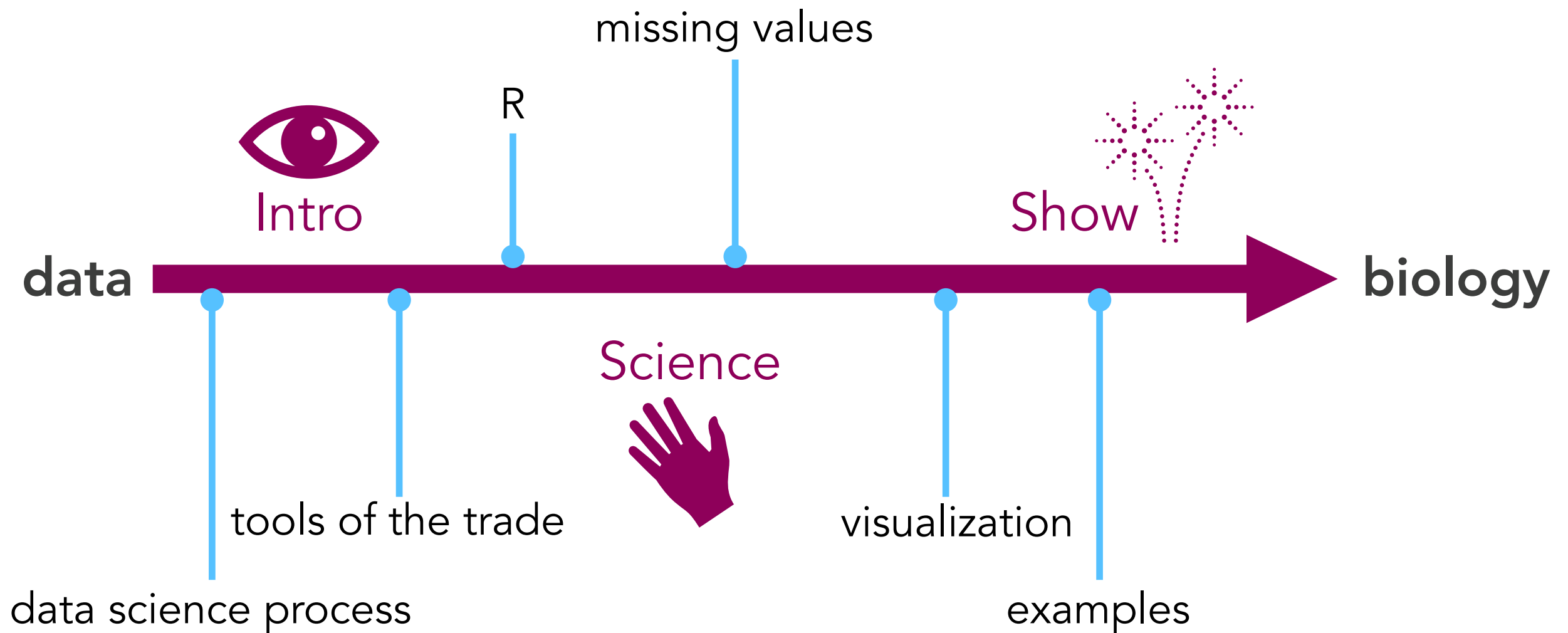
Data science

Data visualization

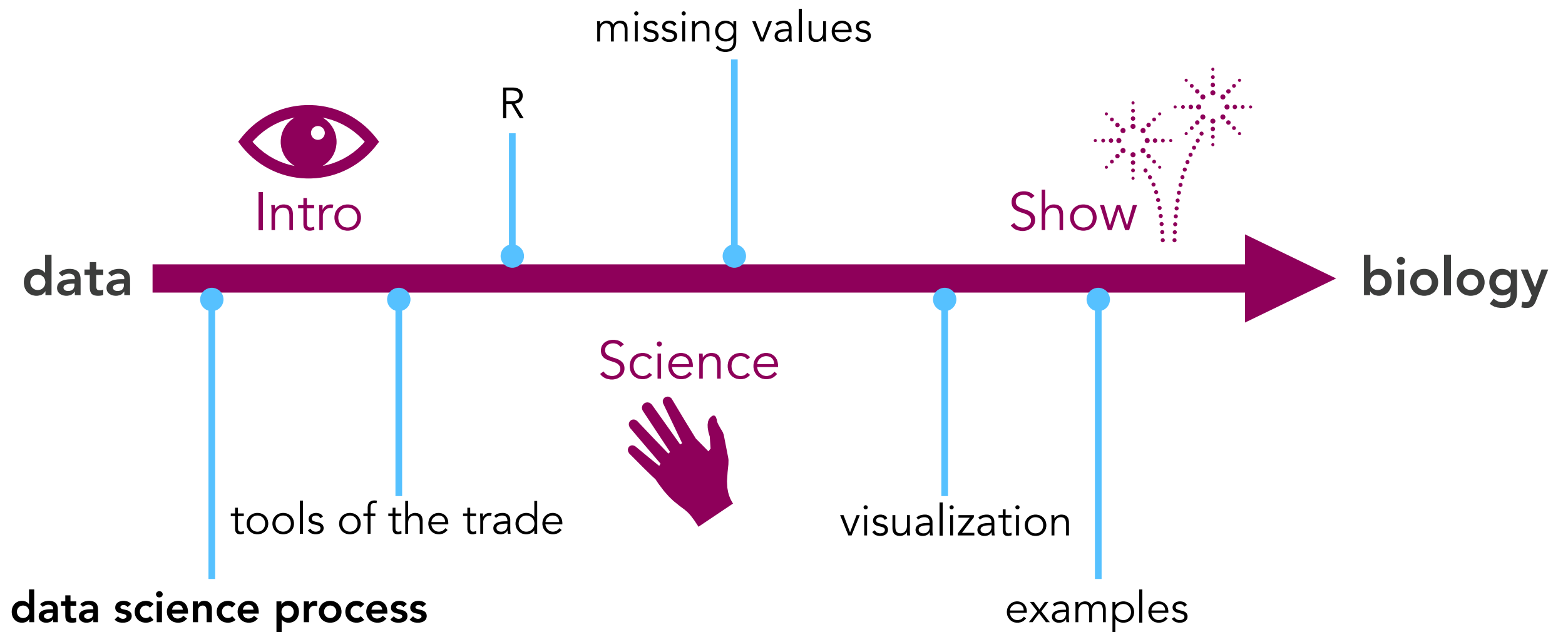
From data to biology

Setting the stage

for today



<https://github.com/higsch/bioinfo-workshop>



Why bioinformatics?

Data

e.g. from an experiment

We intuitively make inferences.

Statistics

A tool to formalize inferences within a mathematical framework.

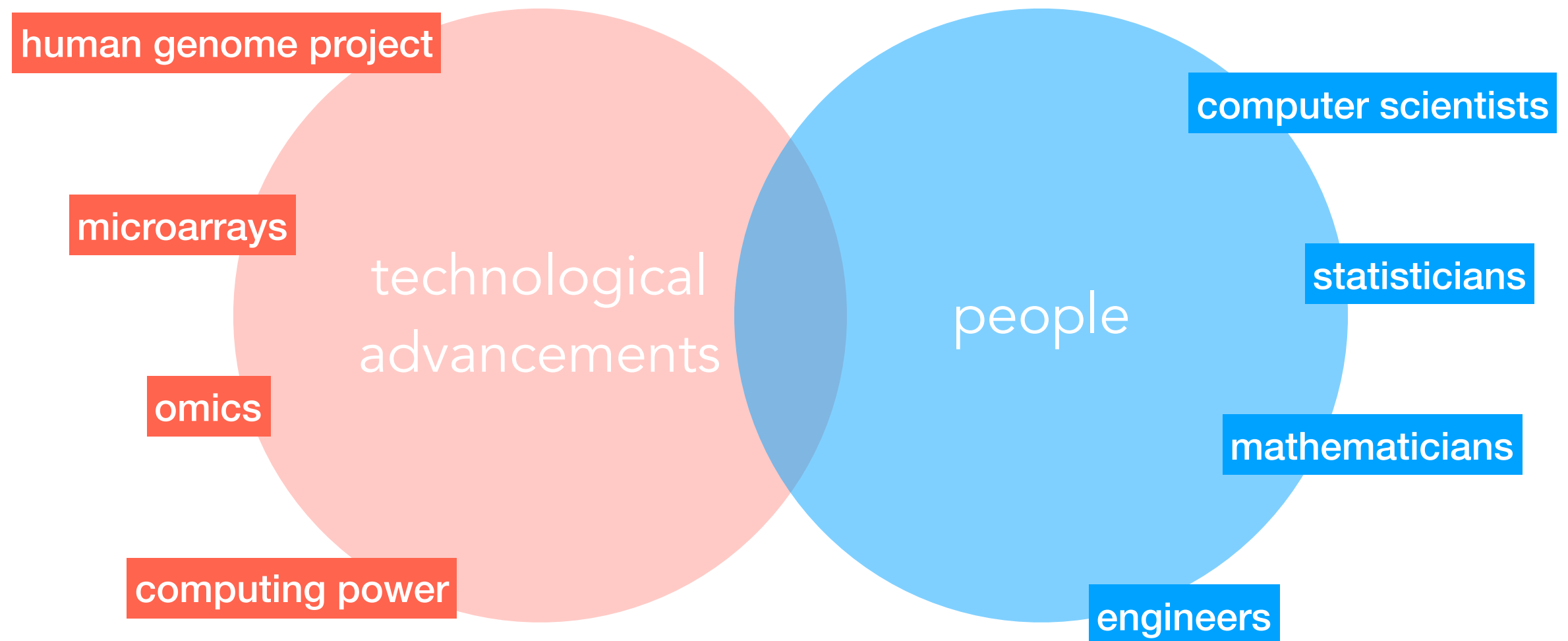
Bioinformatics

Statistics at scale.

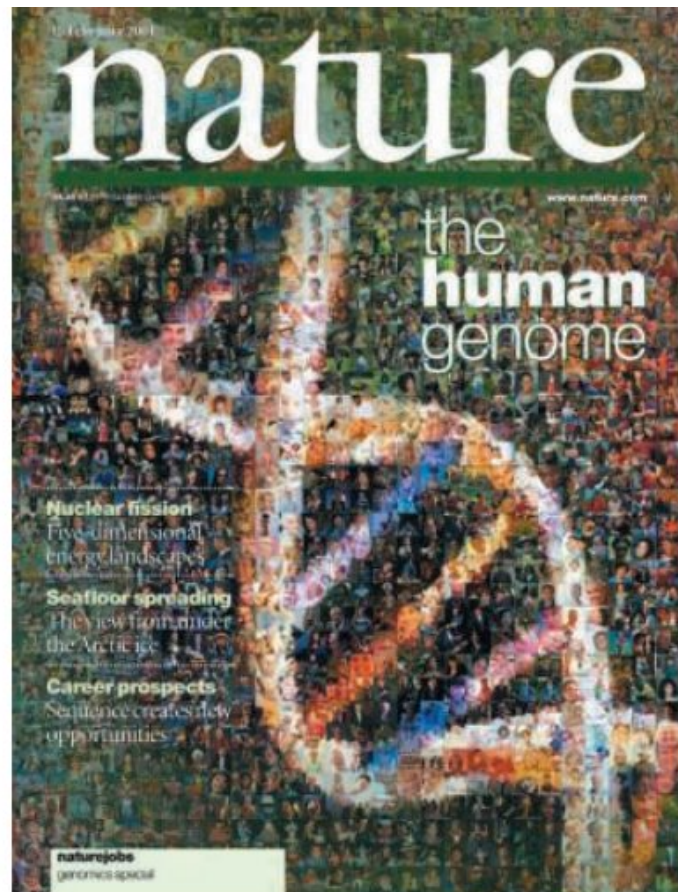
“People don’t want data, they want answers.

— David Hand, Prof. em. Mathematics, Imperial College, London

Where does bioinformatics come from?



Genomics drove the field



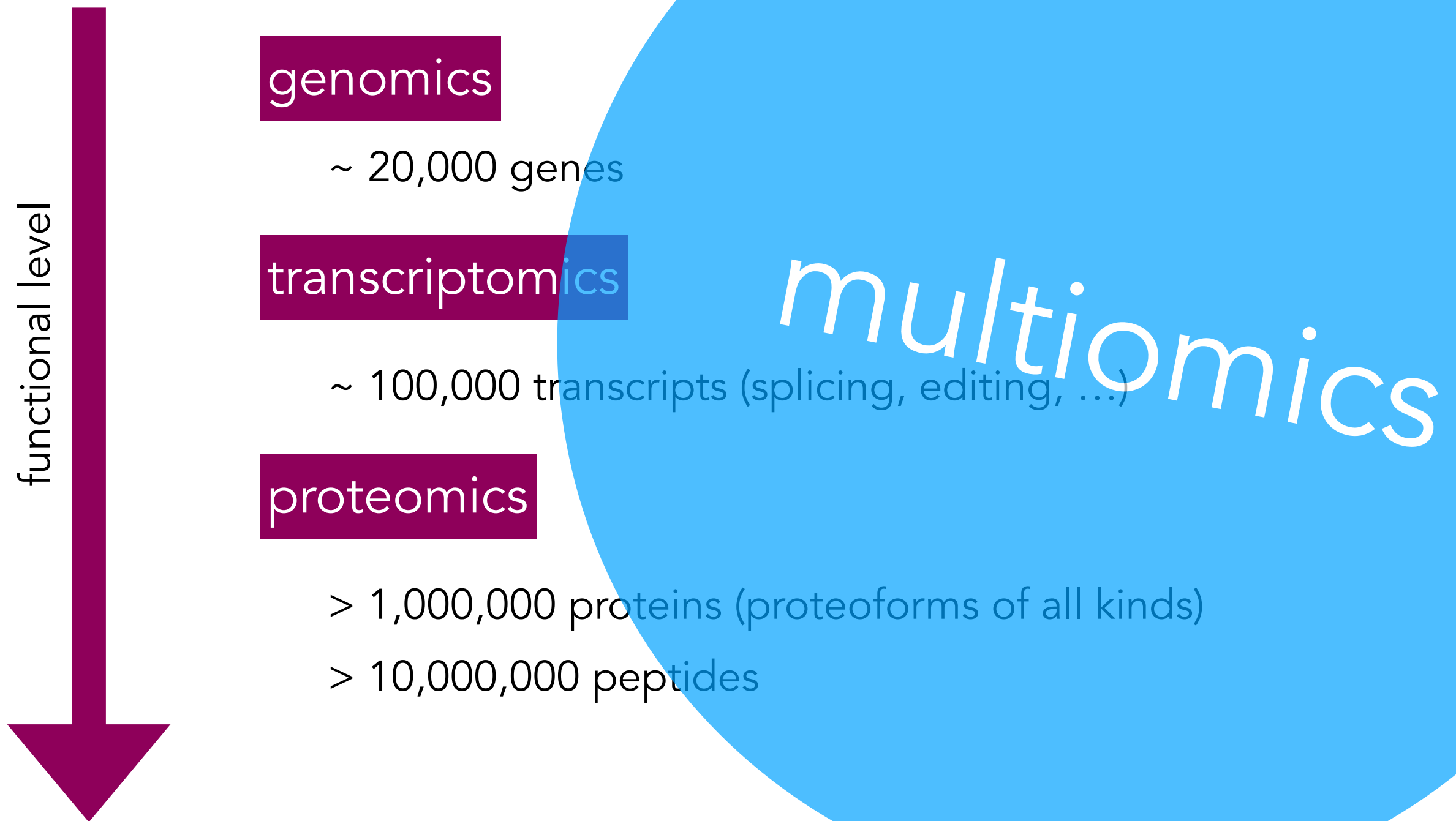
15/02 – 2001



16/02 – 2001

Proteomics can learn a lot from genomics!

We live in the era of big data

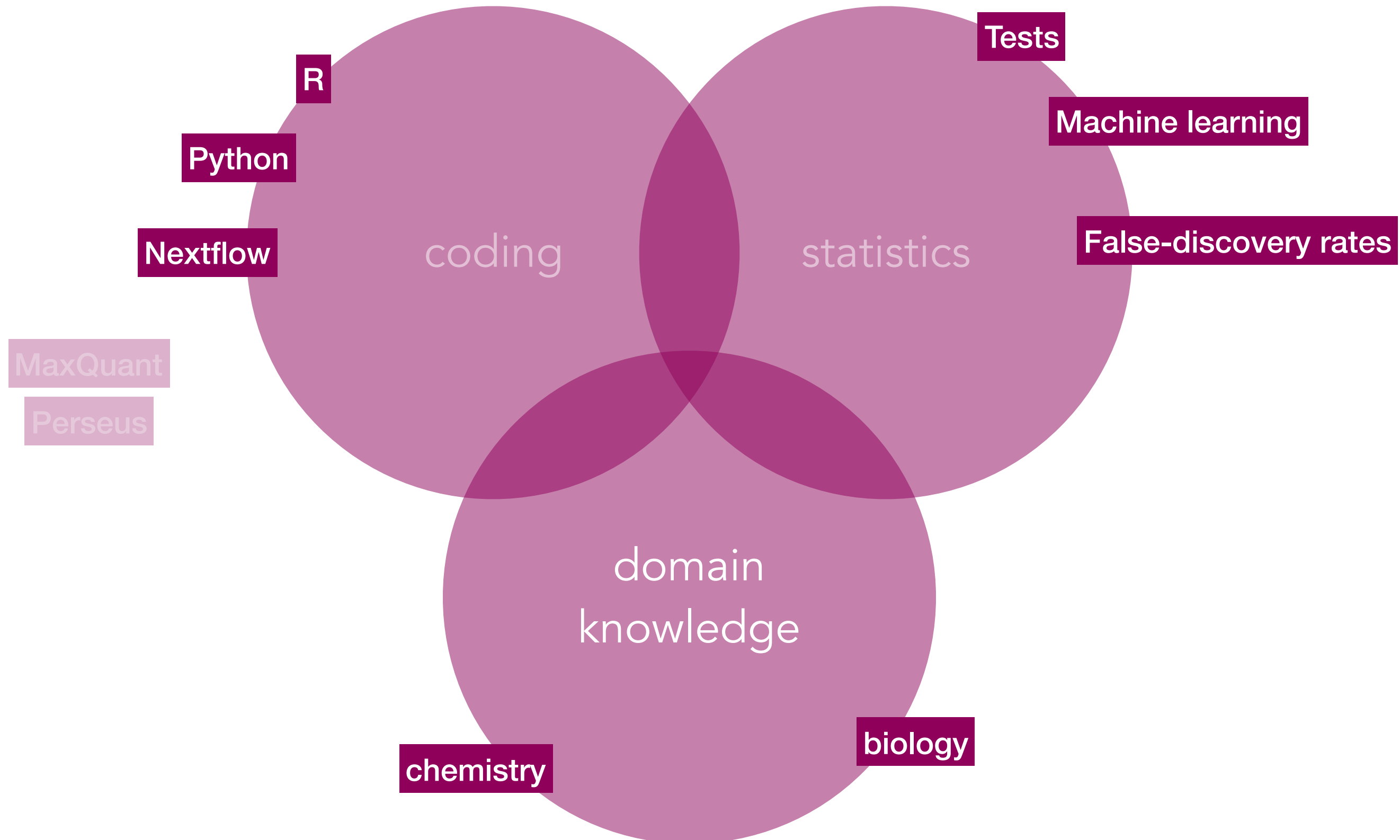


You are all
bioinformaticians

But are you sure about what you are doing?

Please answer a few questions.

What is bioinformatics?



Bioinformatics is a *toolset*

but also a mindset

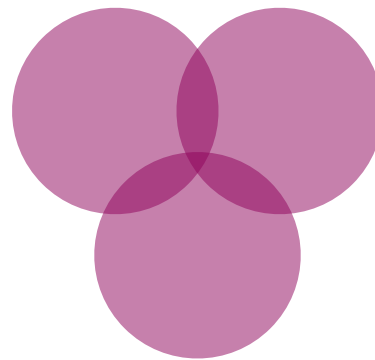
Bioinformatics is...

... not statistics.

... not computer programming.

... not biology.

It is biology *at scale*.



And it has a big problem

ideal case



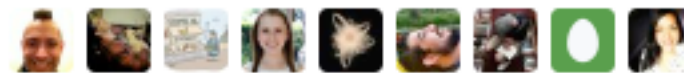
Josh Wills
@josh_wills

 Follow

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

RETWEETS
1,326

LIKES
789



9:55 AM - 3 May 2012

coding

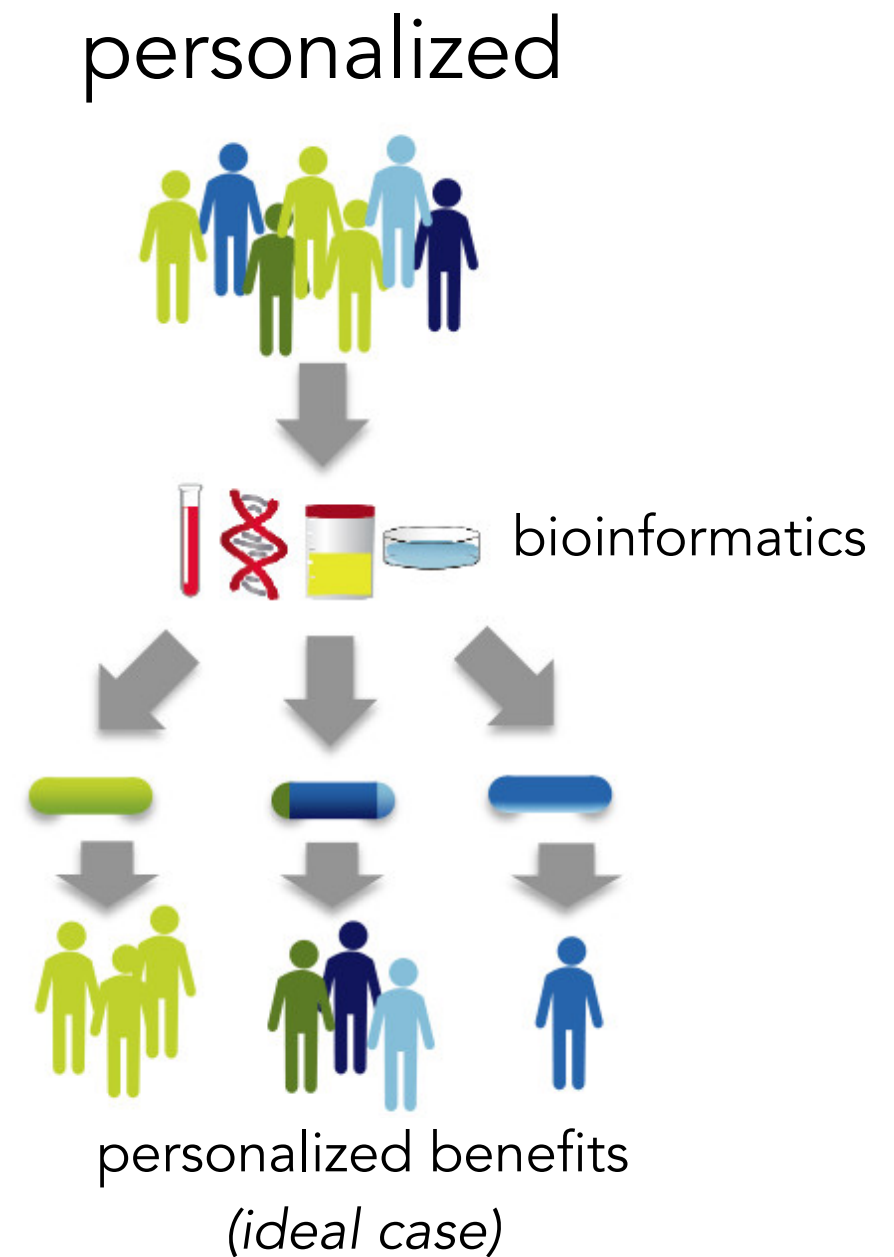
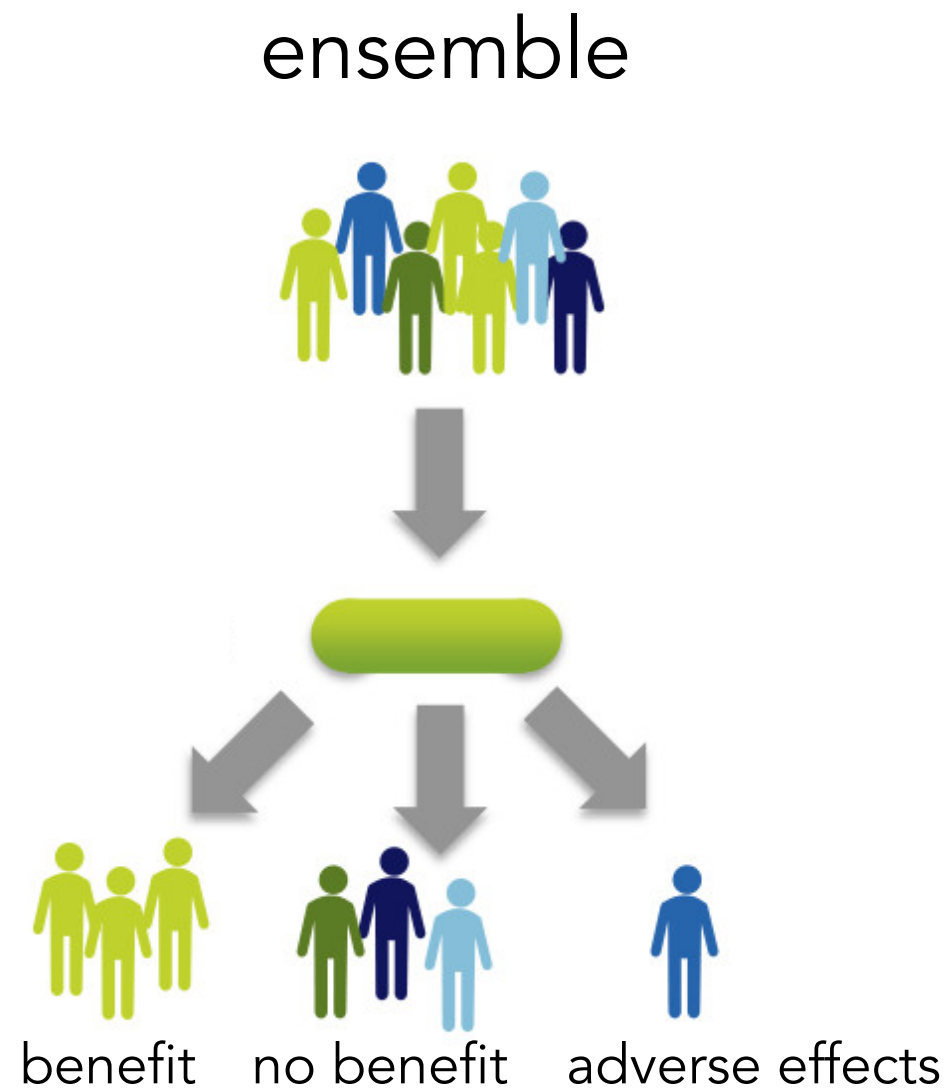
statistics

domain
knowledge

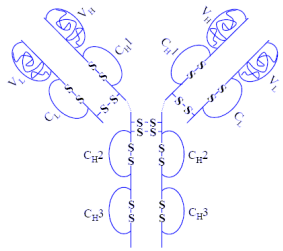
Let's bridge the gap!

♥ Motivation, please! ♥

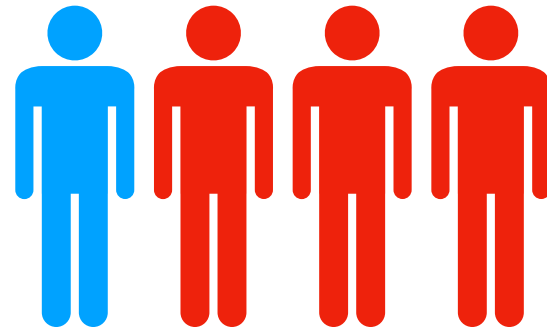
Personalized medicine



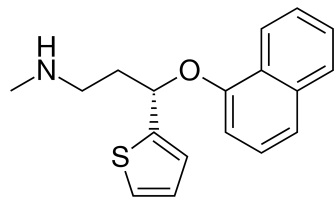
We urgently need personalized medicine!



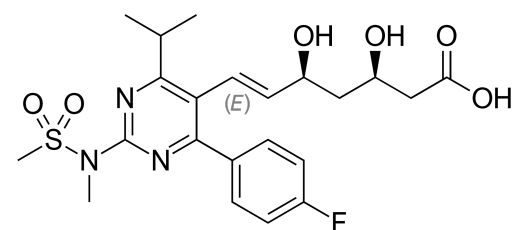
Humira
Arthritis



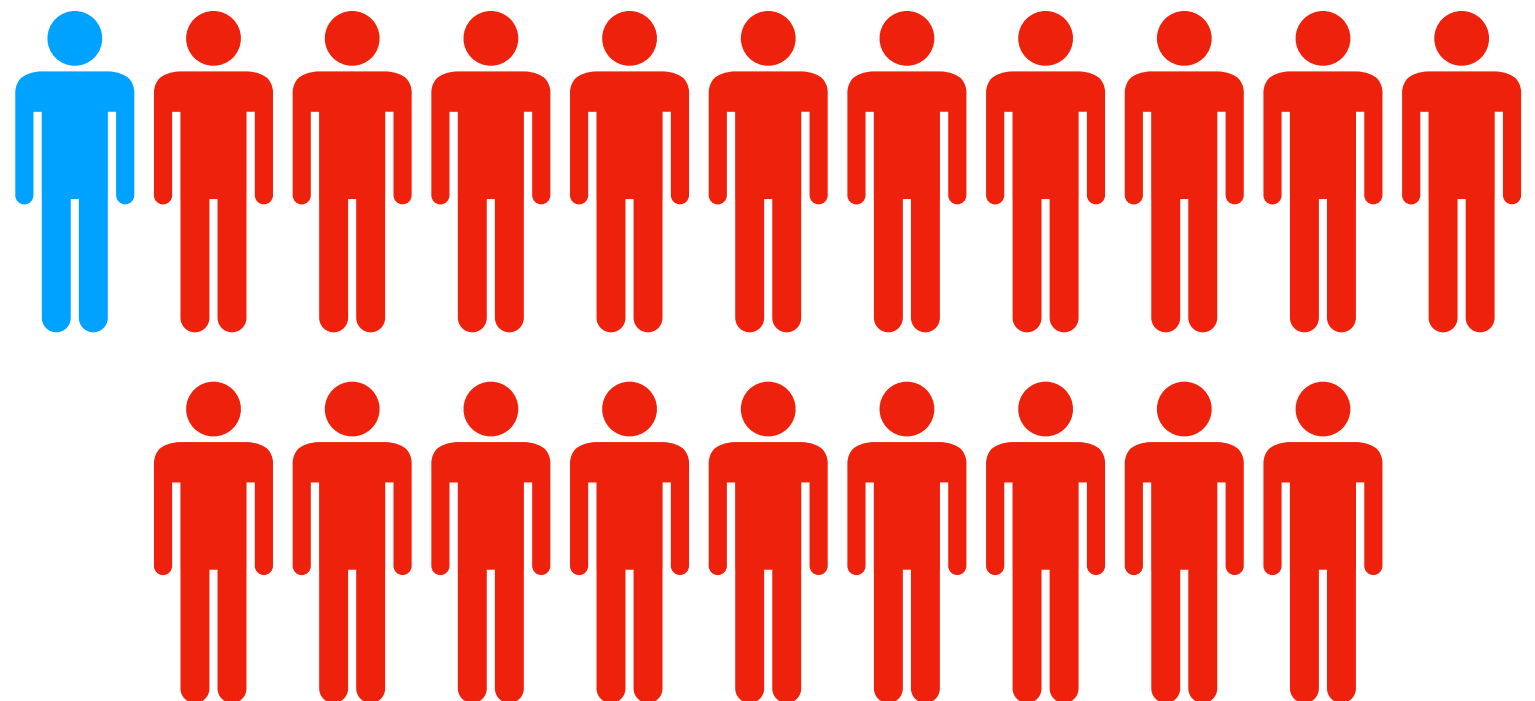
NNT (number needed to treat)



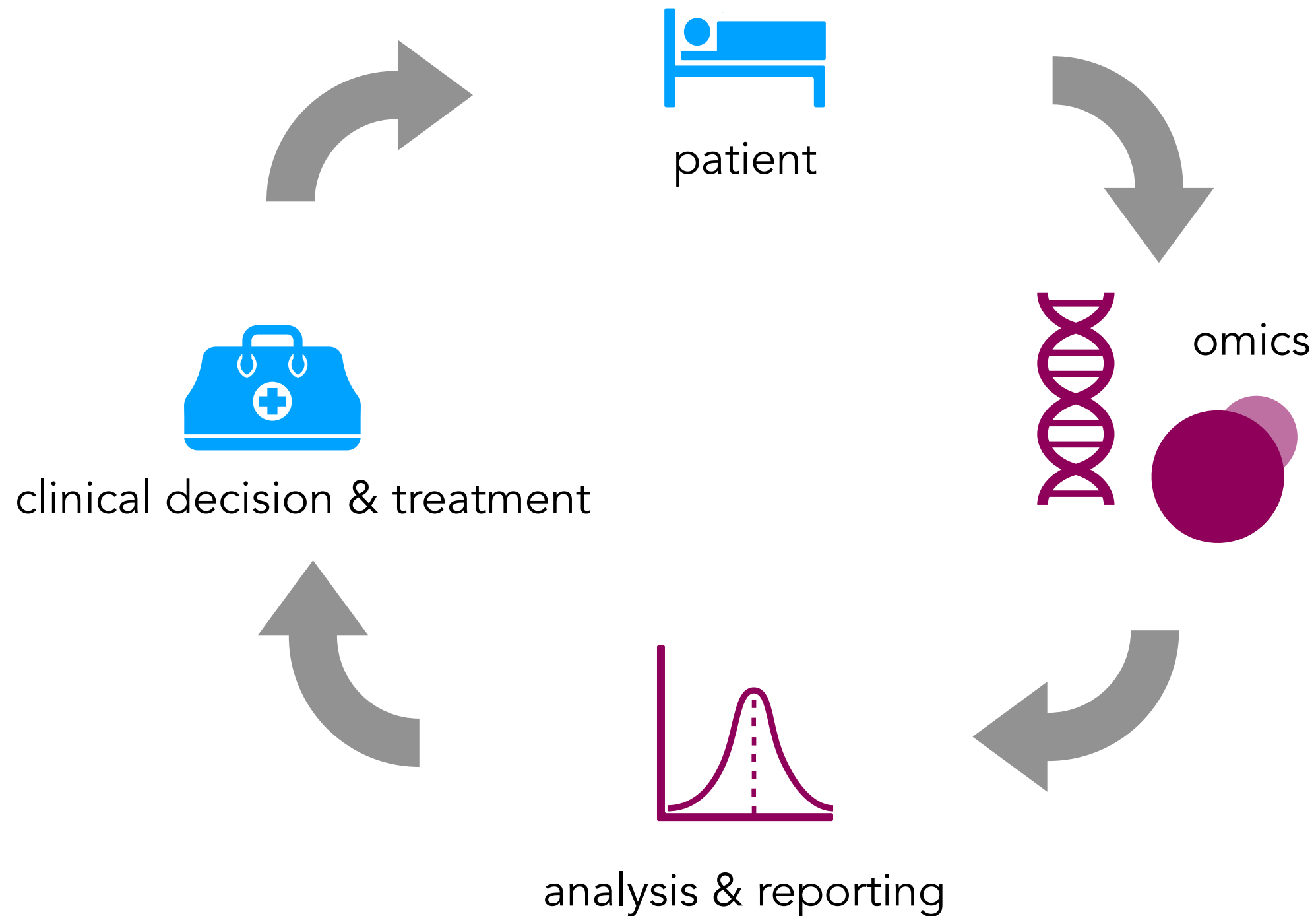
Cymbalta
Depression

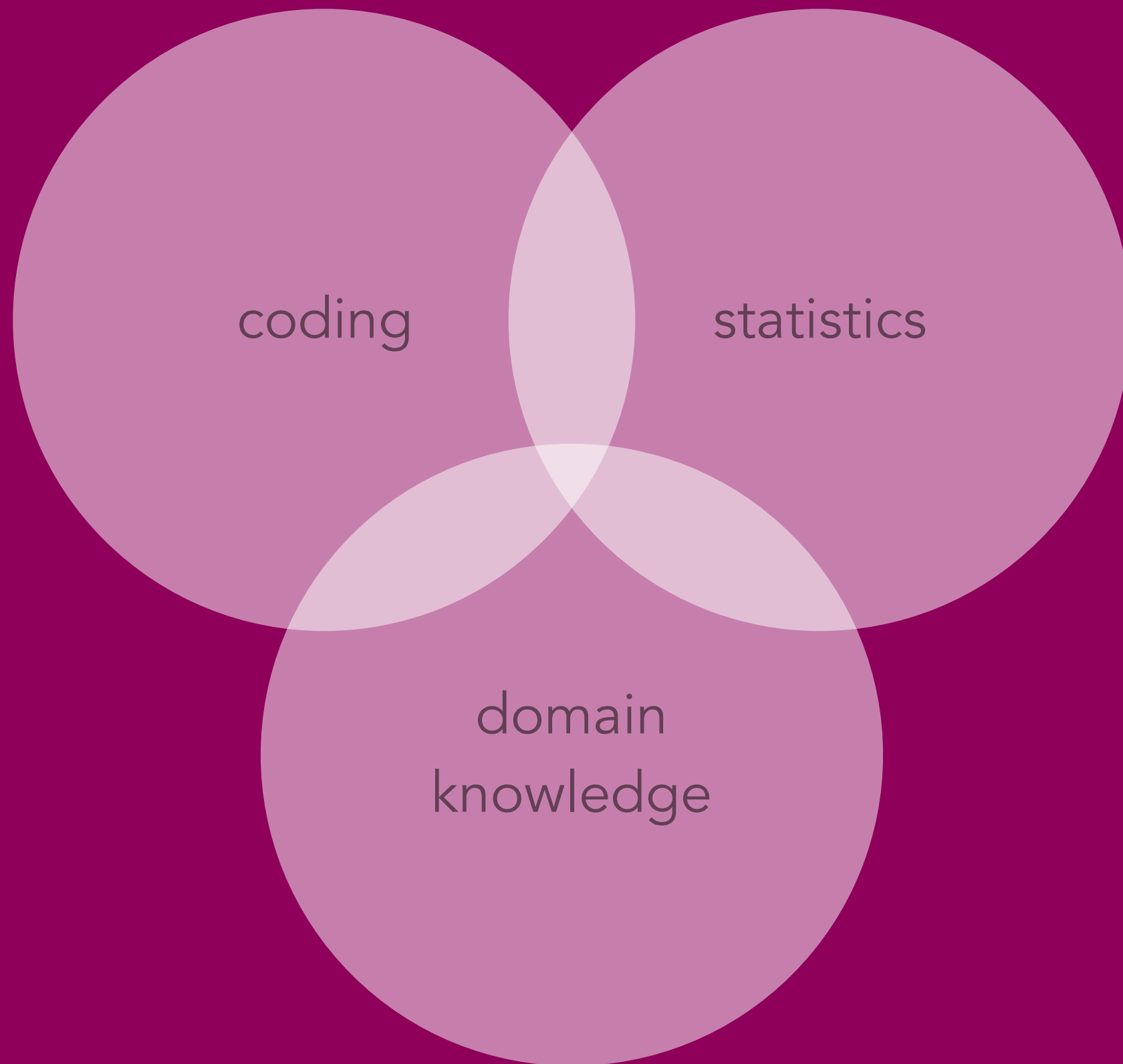


Crestor
High cholesterol




Molecular Tumor Boards





“Do you want to sell sugar water for the rest of your life?
Or do you want to change the world?”



Excel, Perseus, ...

— Steve Jobs when recruiting Pepsi CEO John Sculley



R, Python, JavaScript, ...