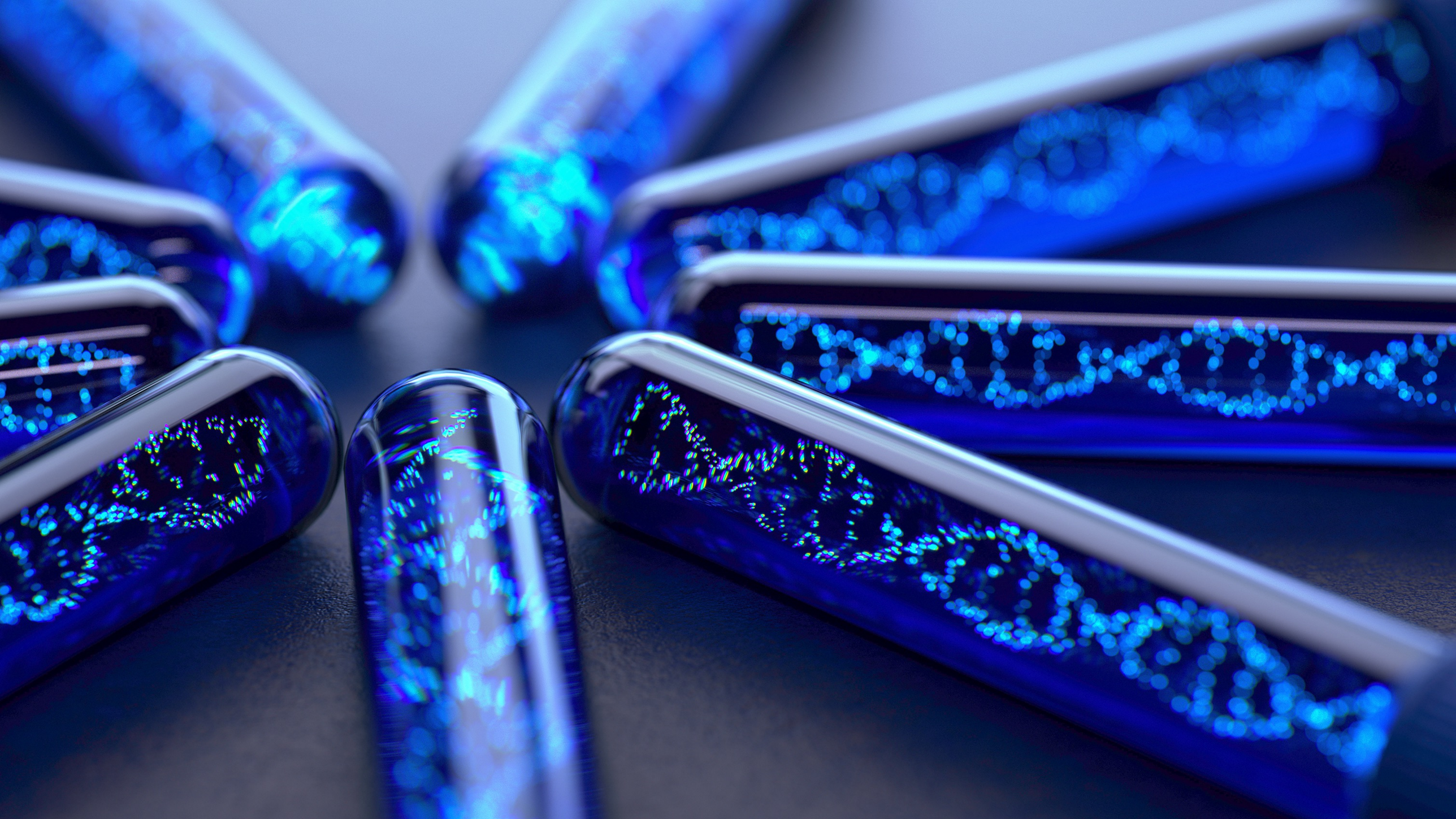
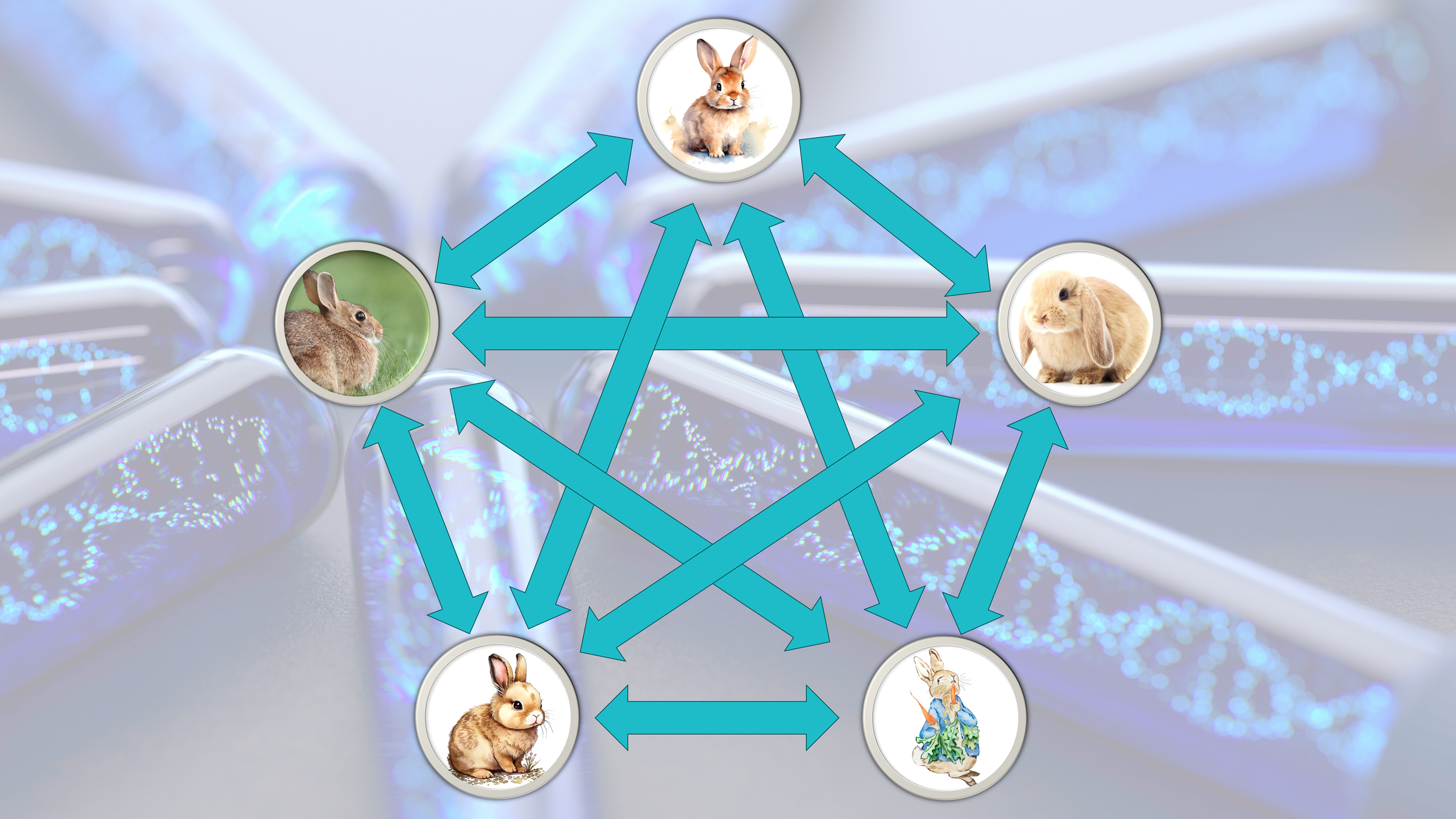
Softwire

Ben Below - Softwire

Chromosomes in the cloud

Building a scalable genomic matching algorithm



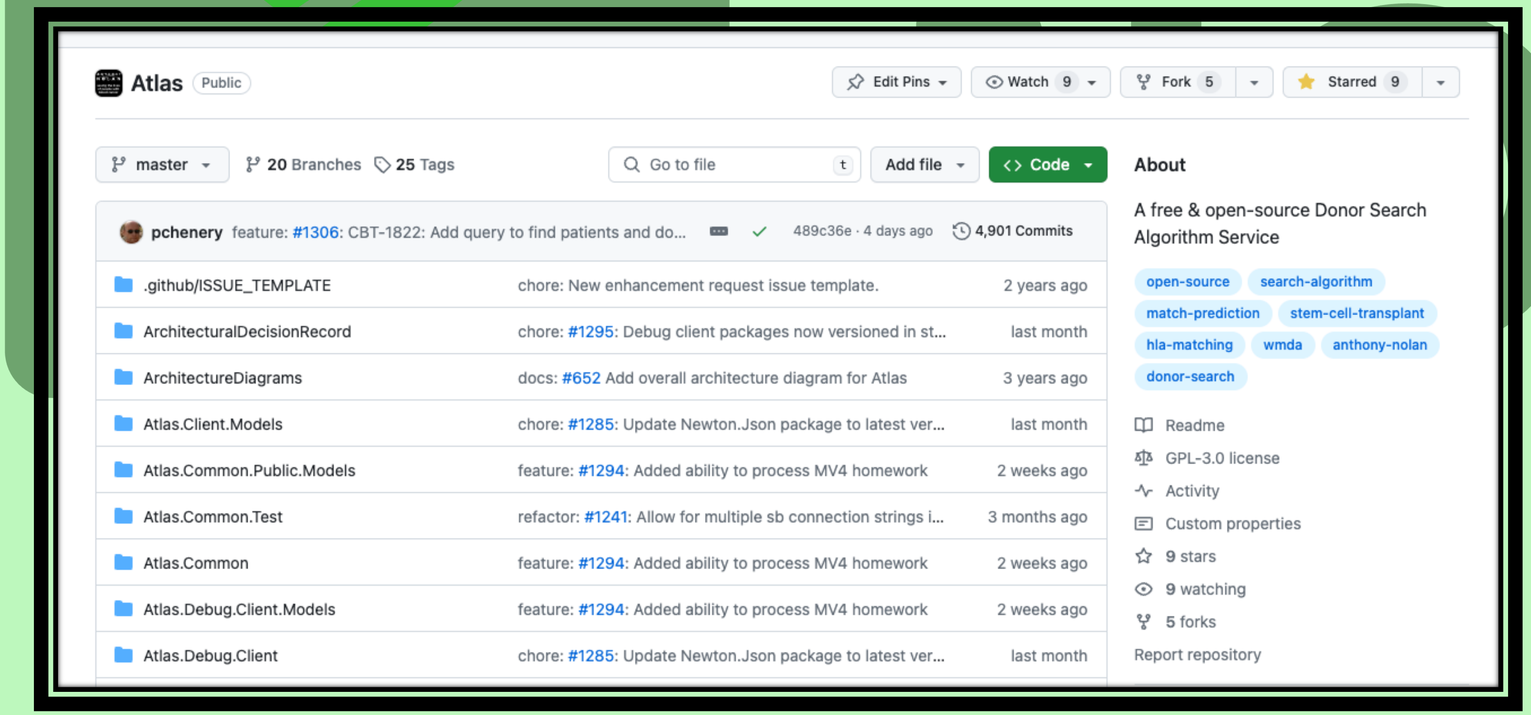


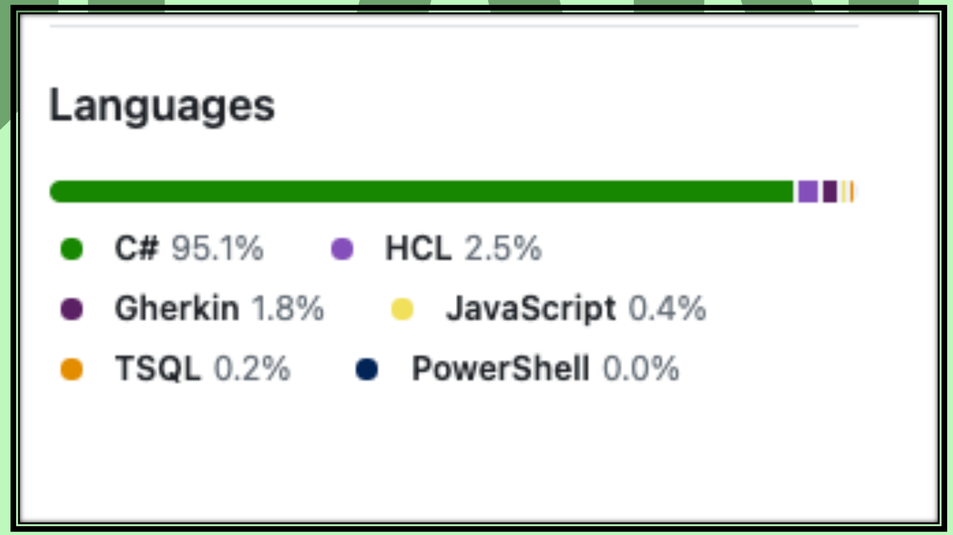

Your immune system can identify you

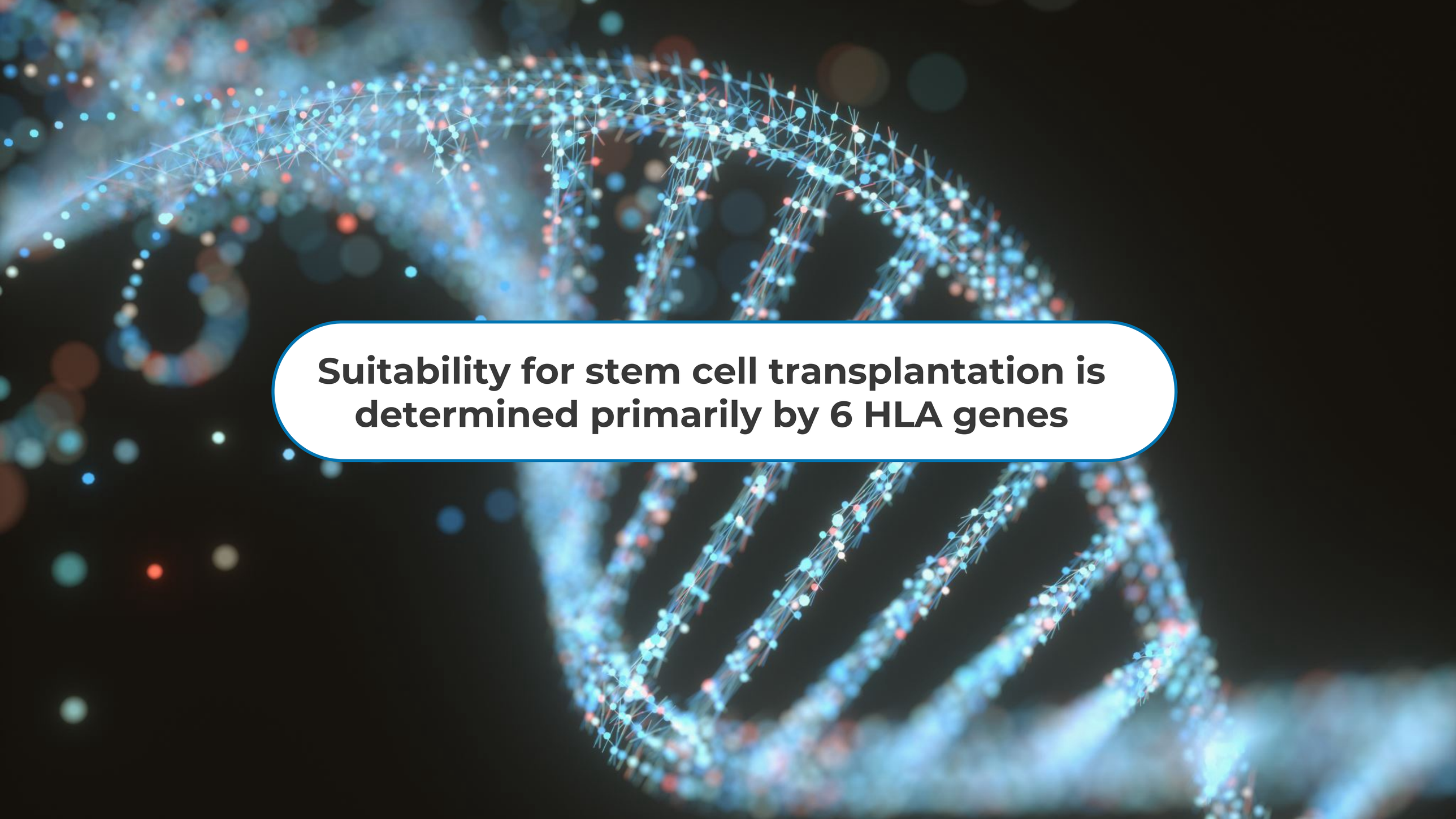
Your immune system can identify you, and you’re unique!

ANTHONY
NOLAN

ANTHONY NOLAN





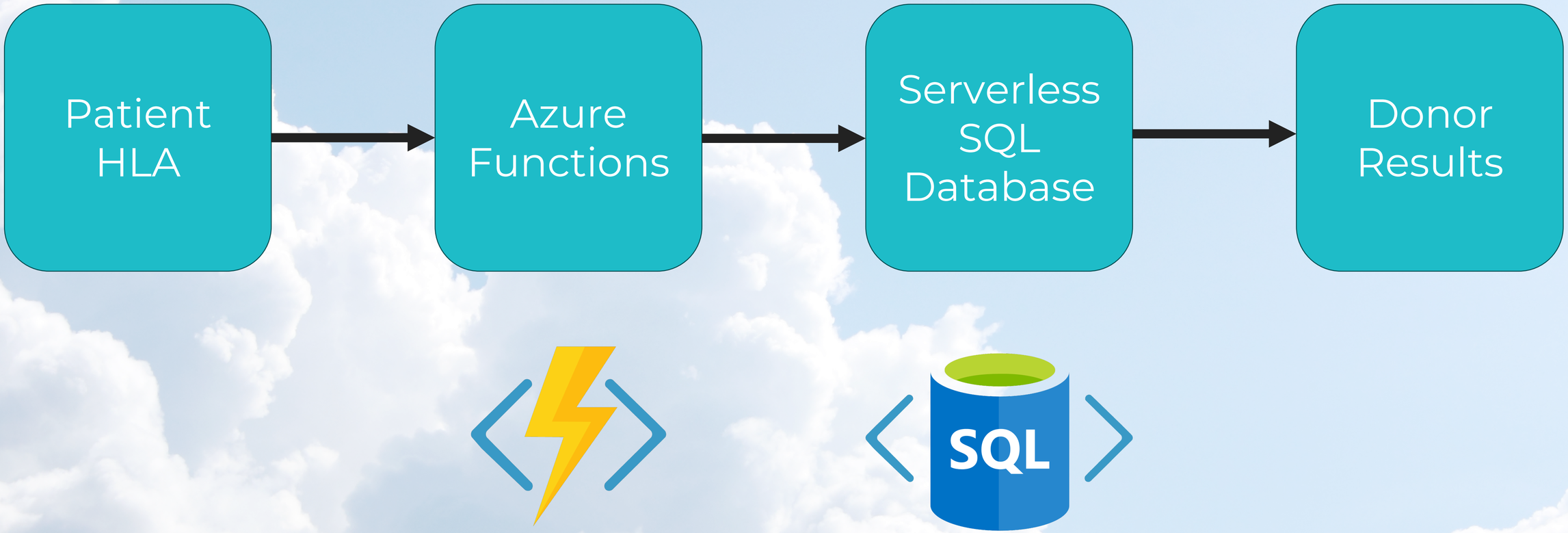

Suitability for stem cell transplantation is determined primarily by 6 HLA genes

We may only know that a donor has one of 500 different options for each gene

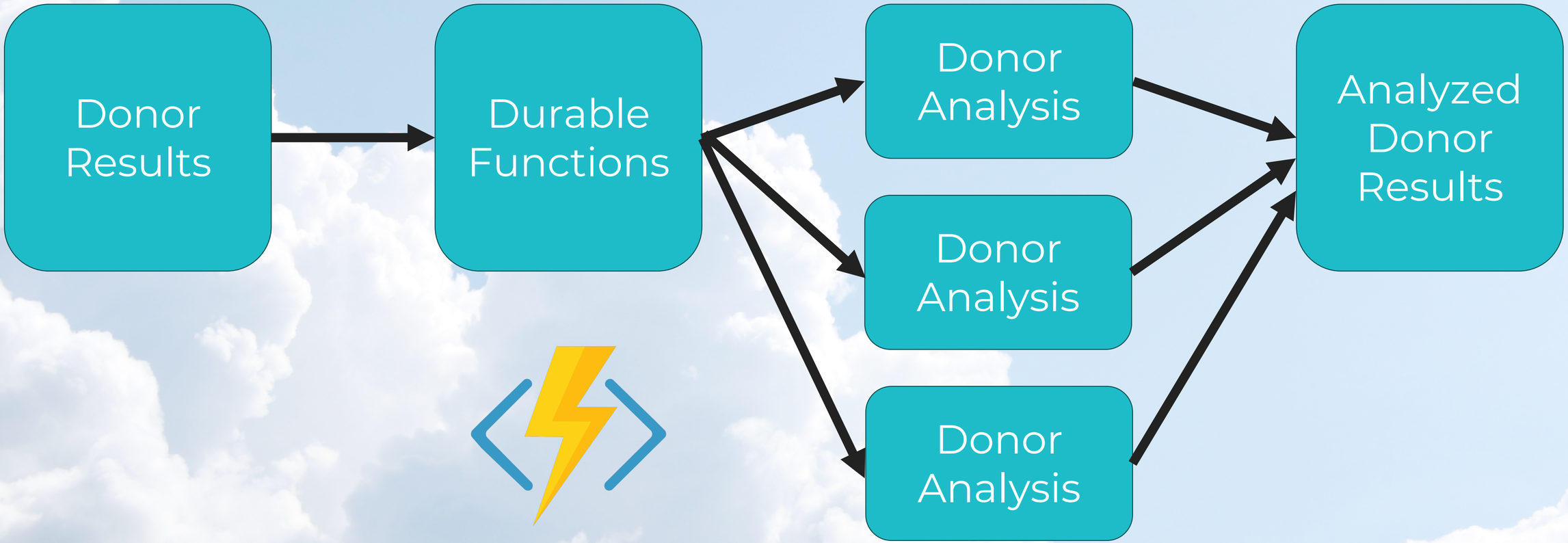
No Ambiguity	Some Ambiguity	Large Ambiguity
A*01:01	A*01:AB	A*01:XX
	A*01:01	A*01:01
	OR	**OR**
	A*01:02	A*01:02
		OR 453 MORE...

We worked closely with Anthony Nolan's scientists to build a scalable algorithm

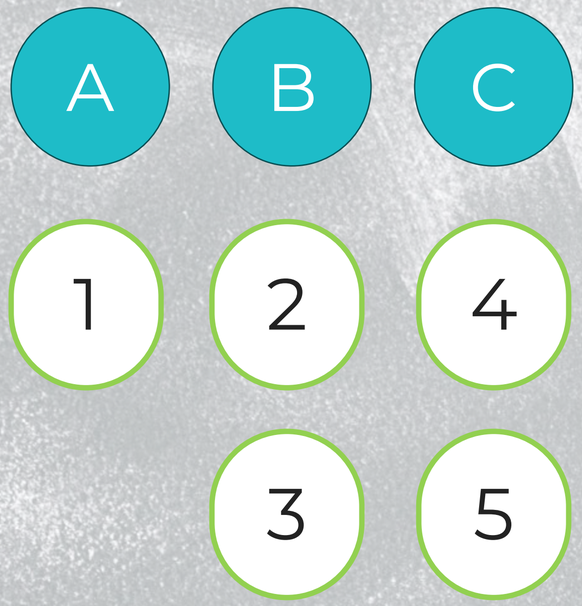
Matching Algorithm

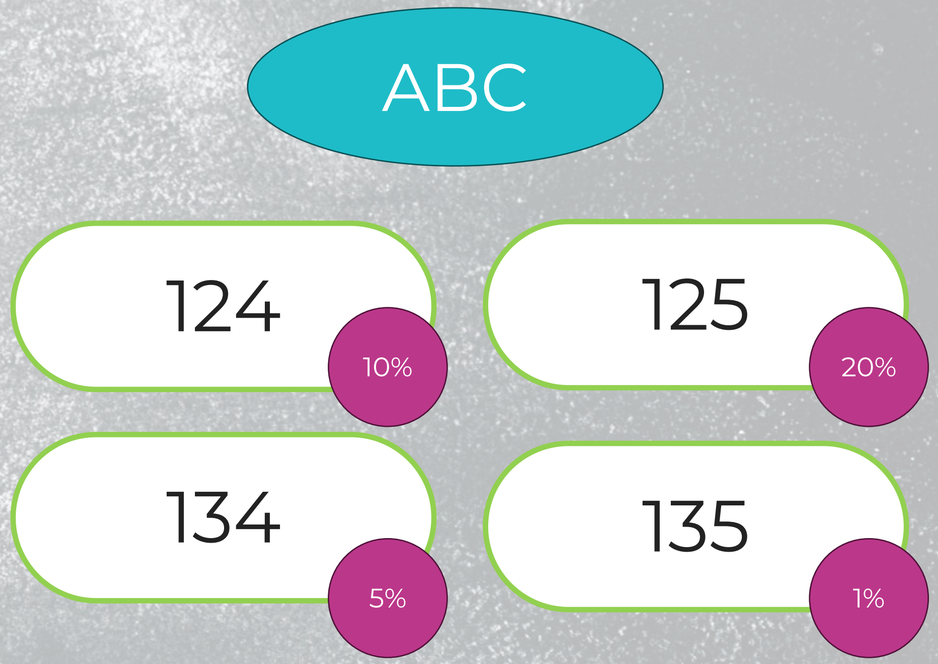


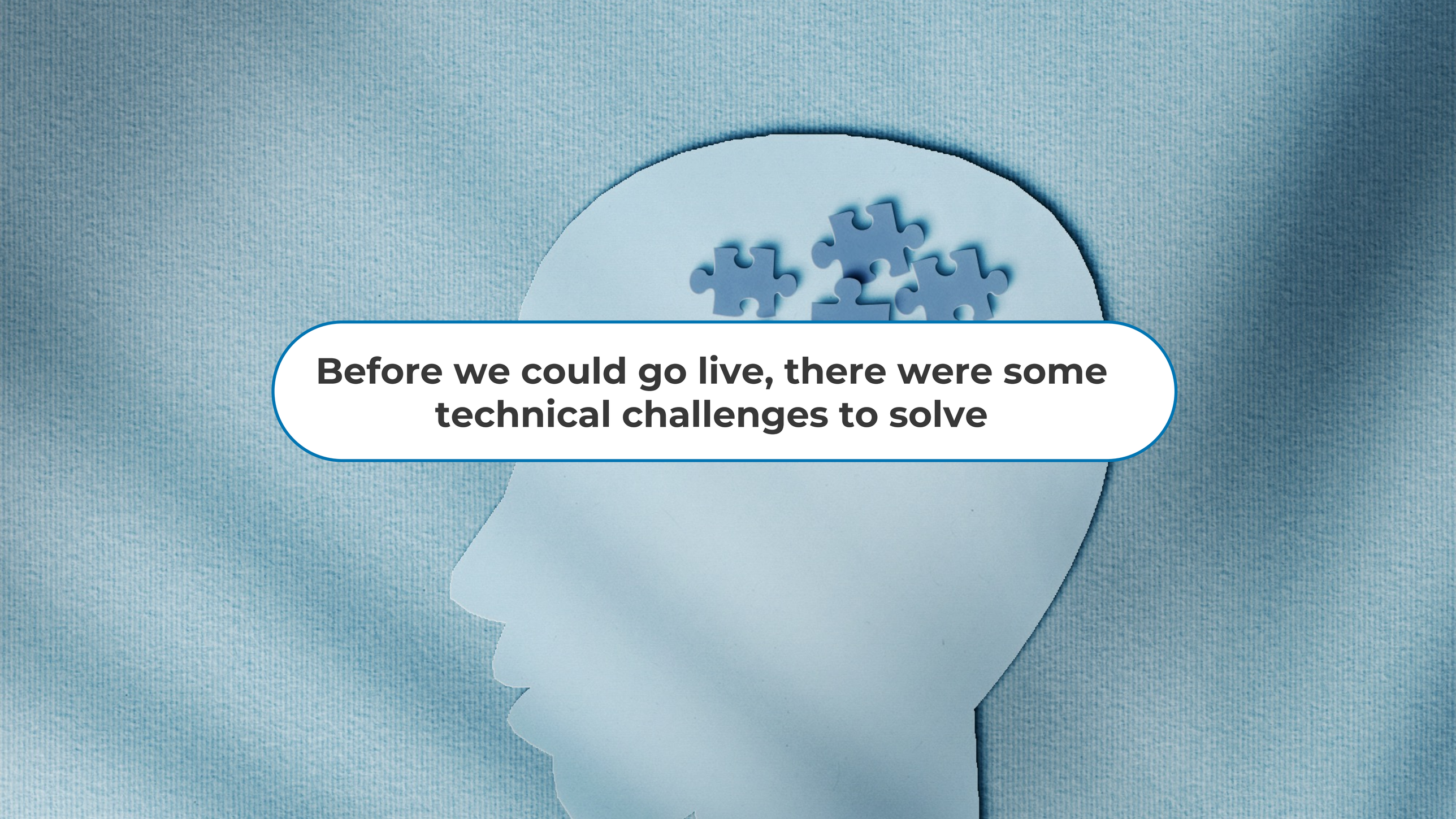
Match Prediction Algorithm

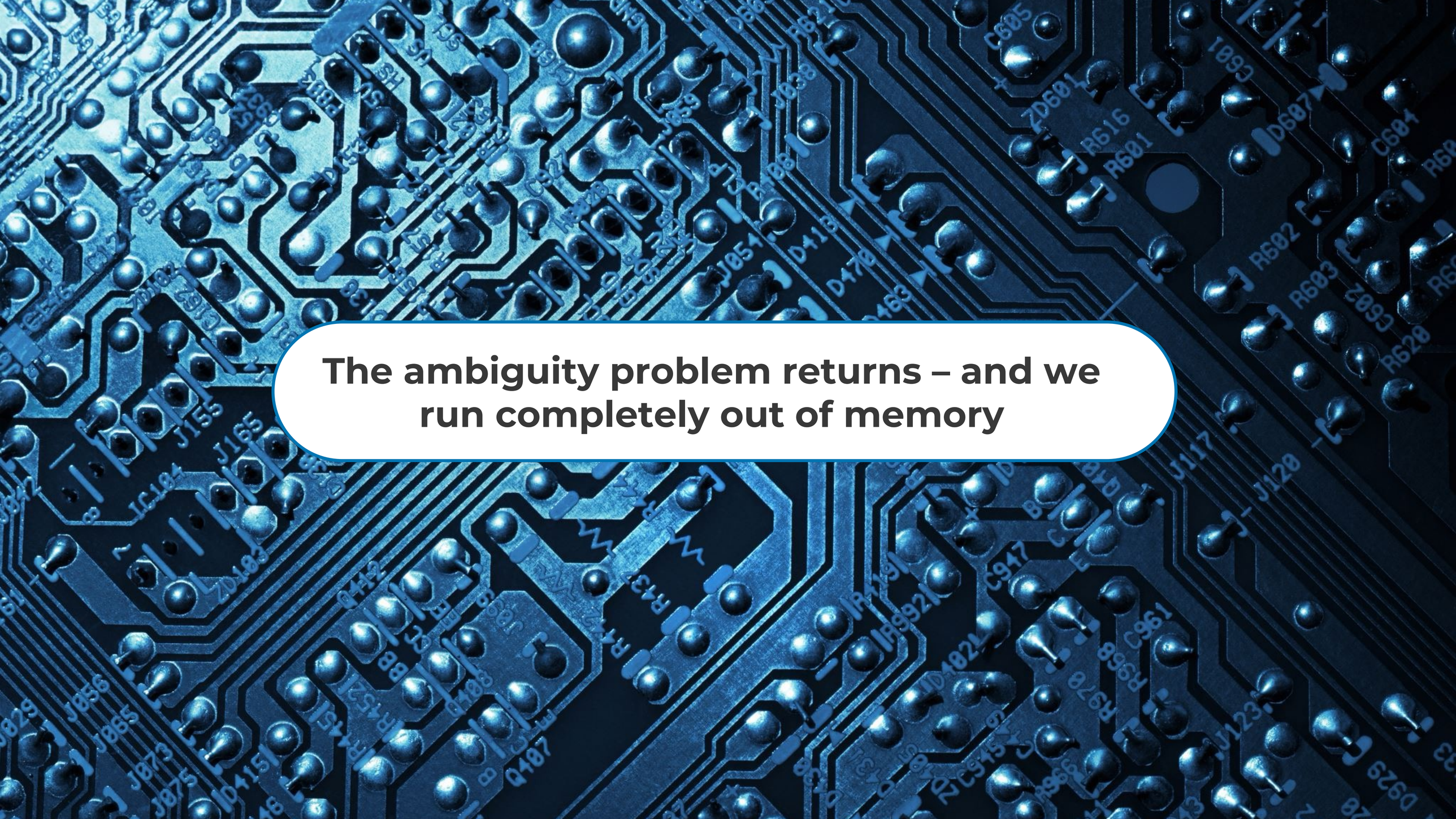


Donor Analysis






Before we could go live, there were some technical challenges to solve


The ambiguity problem returns – and we run completely out of memory


A
B
C

A

B

C

DRB1

DQB1

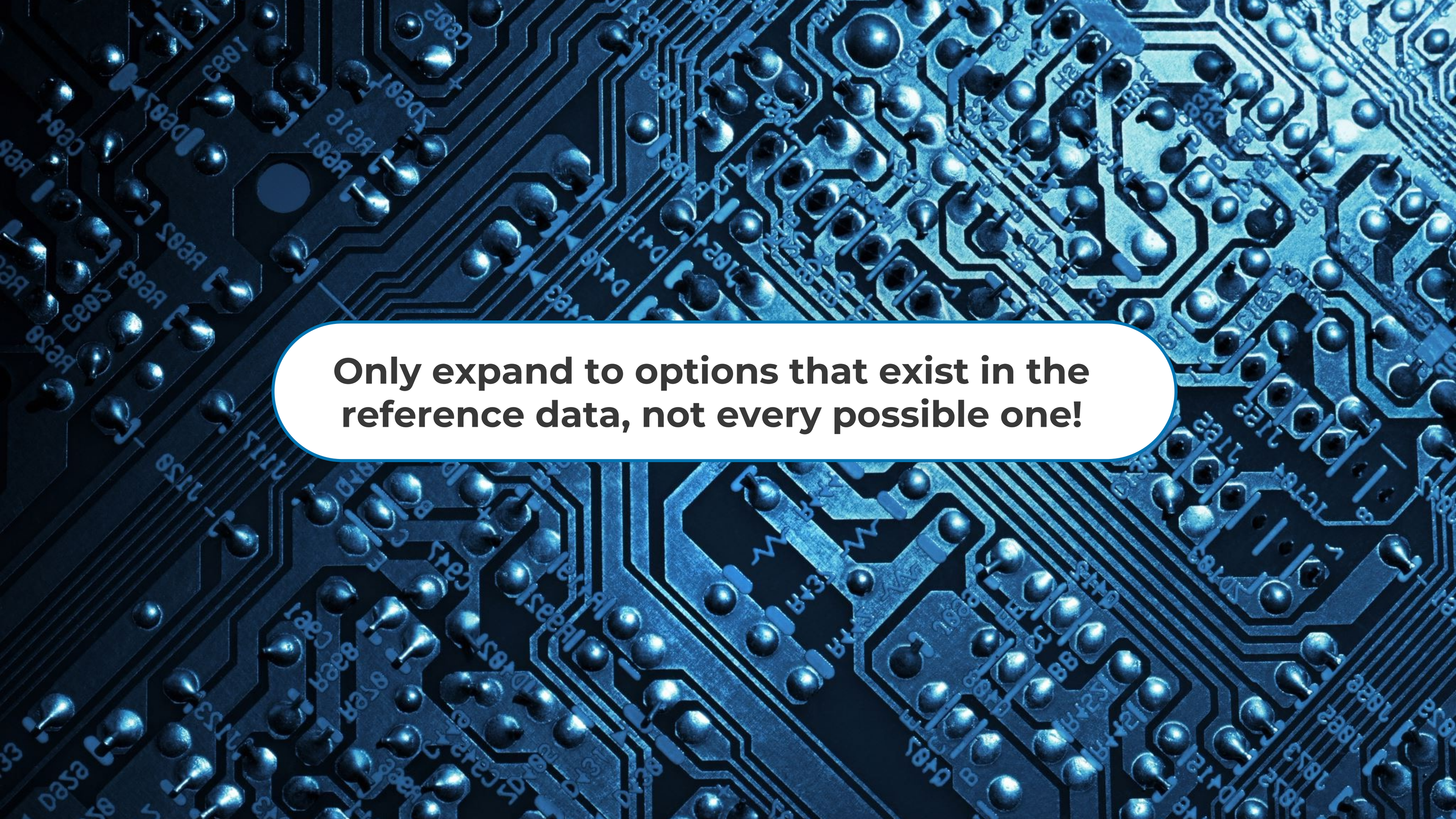
Gigabyte = 10^9

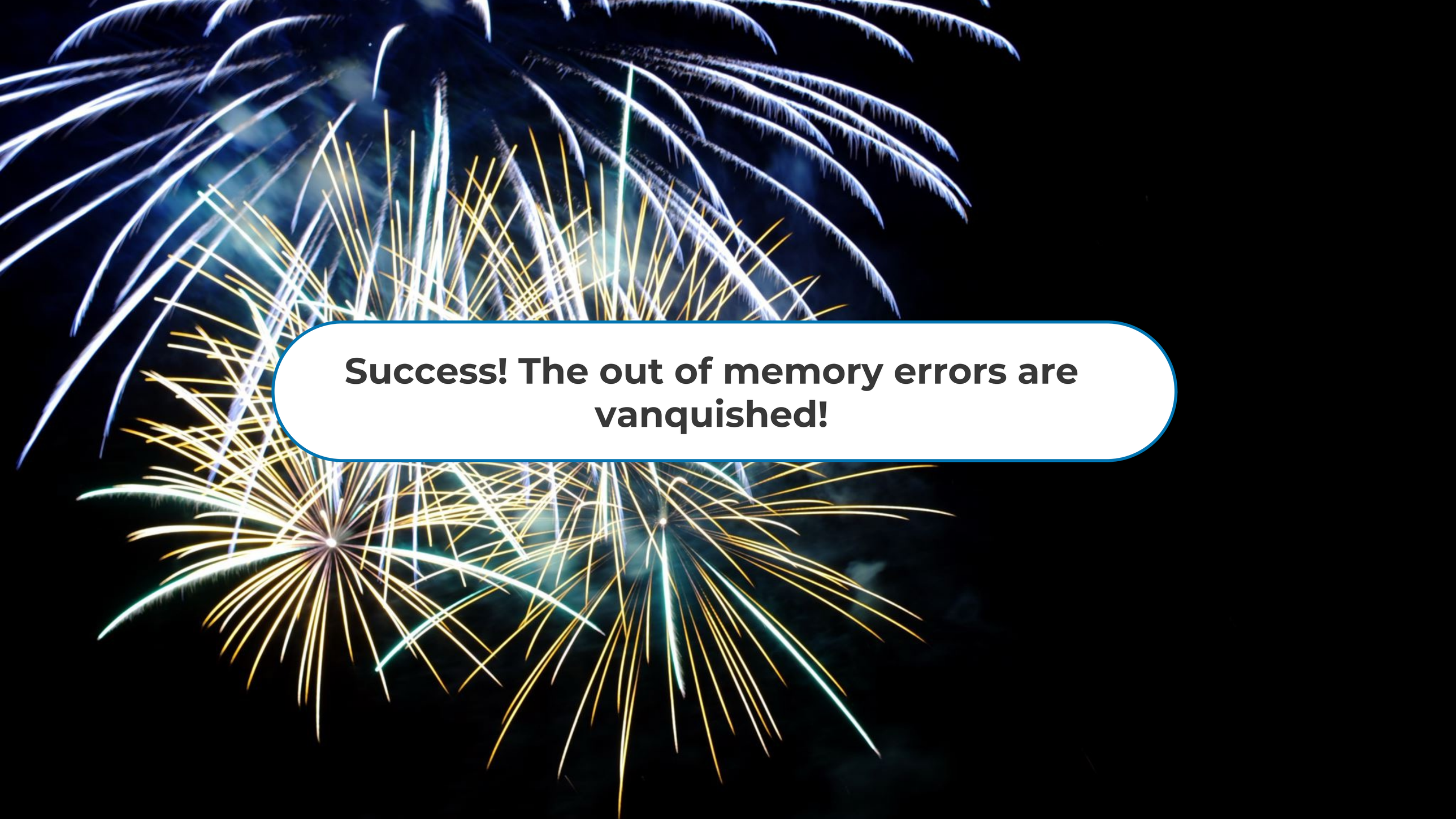
Terabyte = 10^{12}

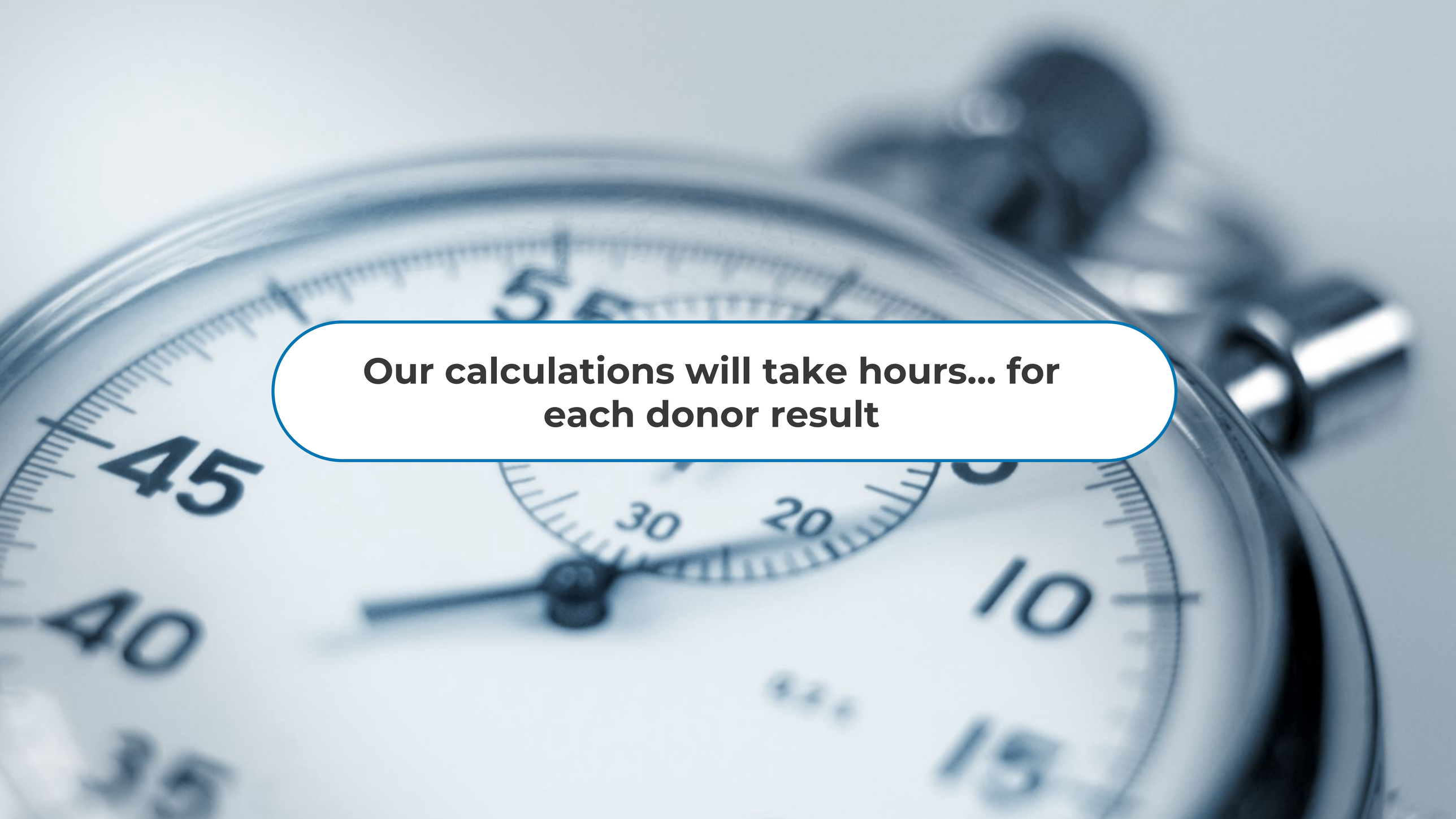
Zettabyte = 10^{21}

Yottabyte = 10^{24}

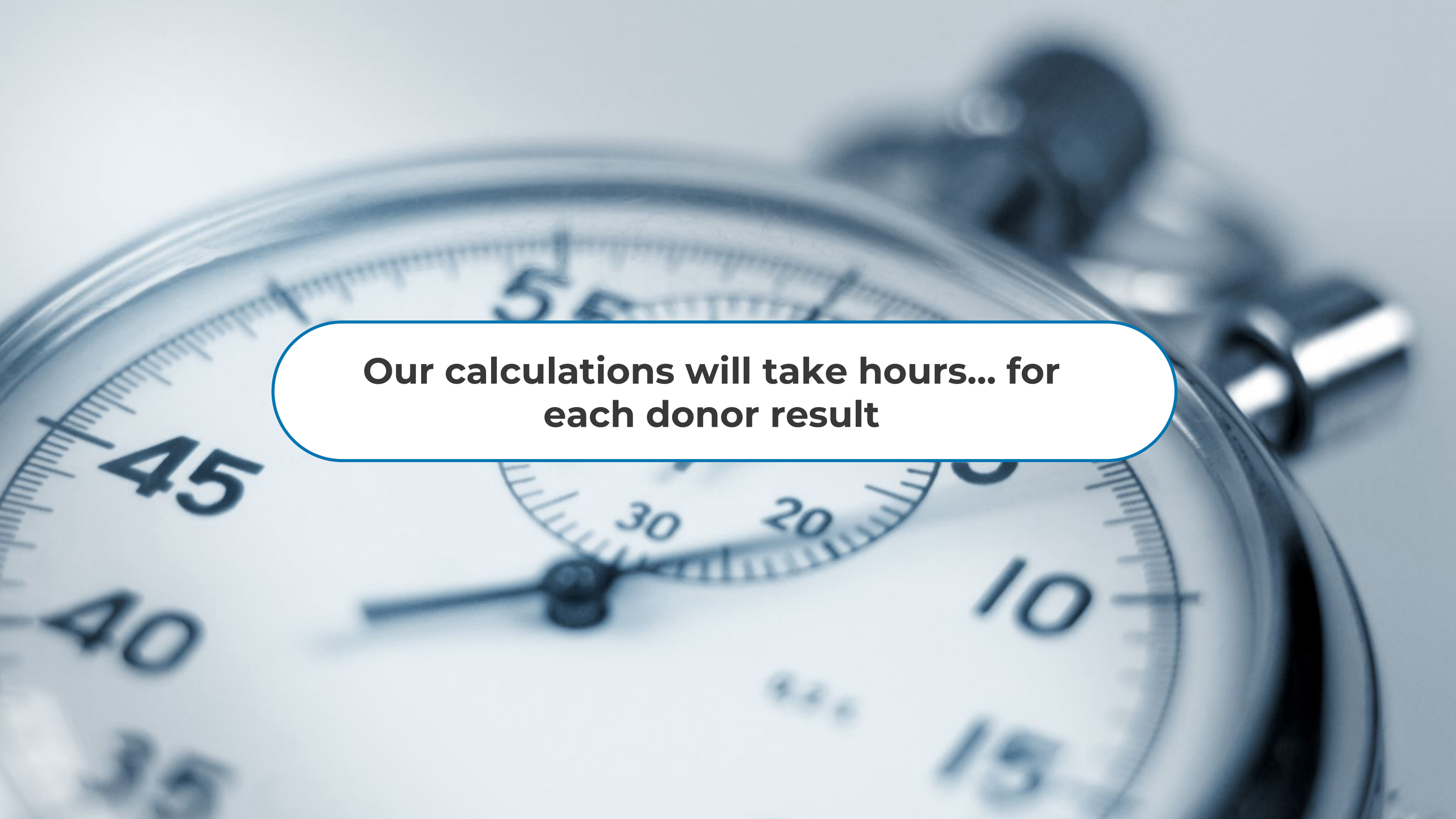

No computer on Earth could hold all possible genetic combinations


Only expand to options that exist in the reference data, not every possible one!


Success! The out of memory errors are vanquished!



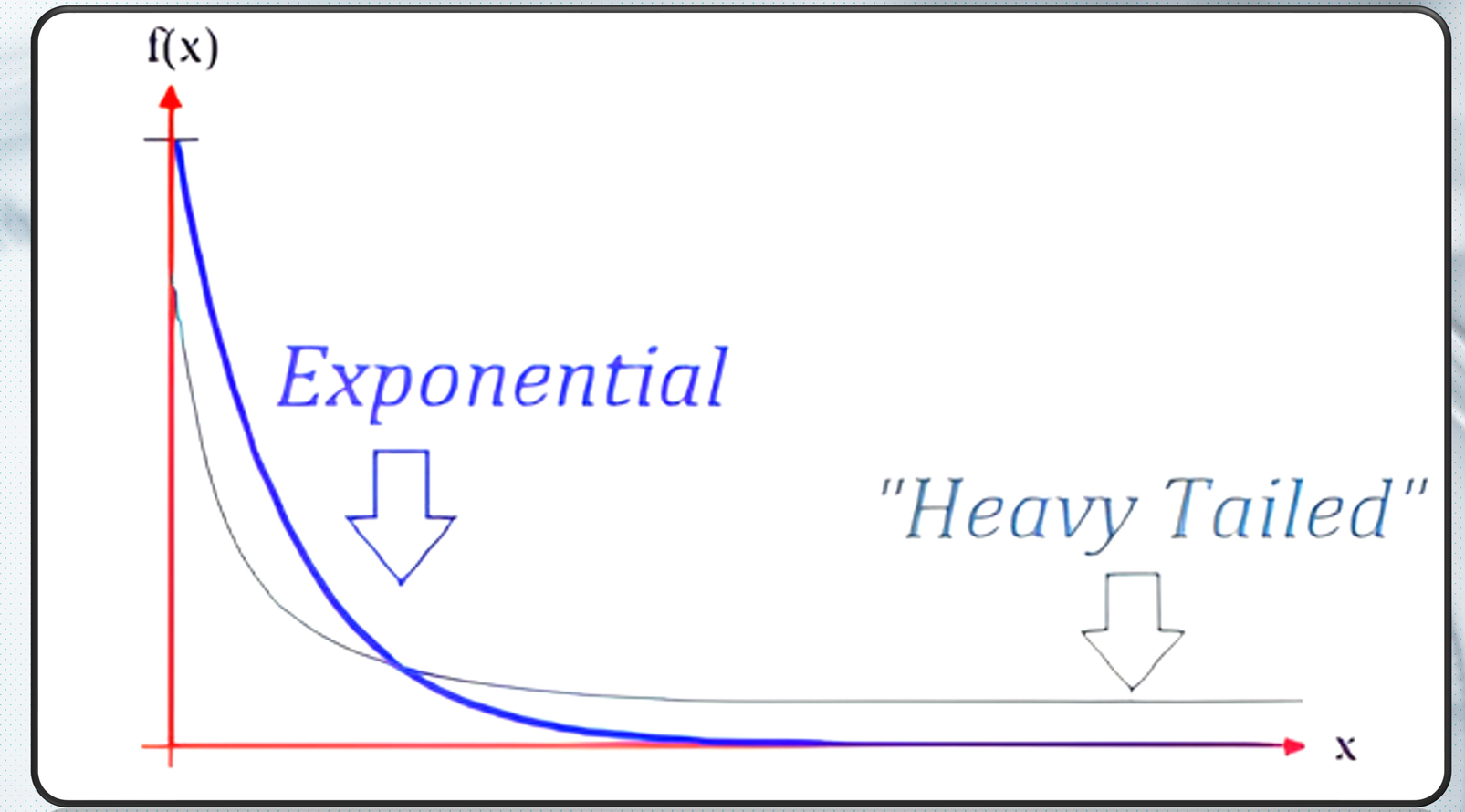
Our calculations will take hours… for each donor result

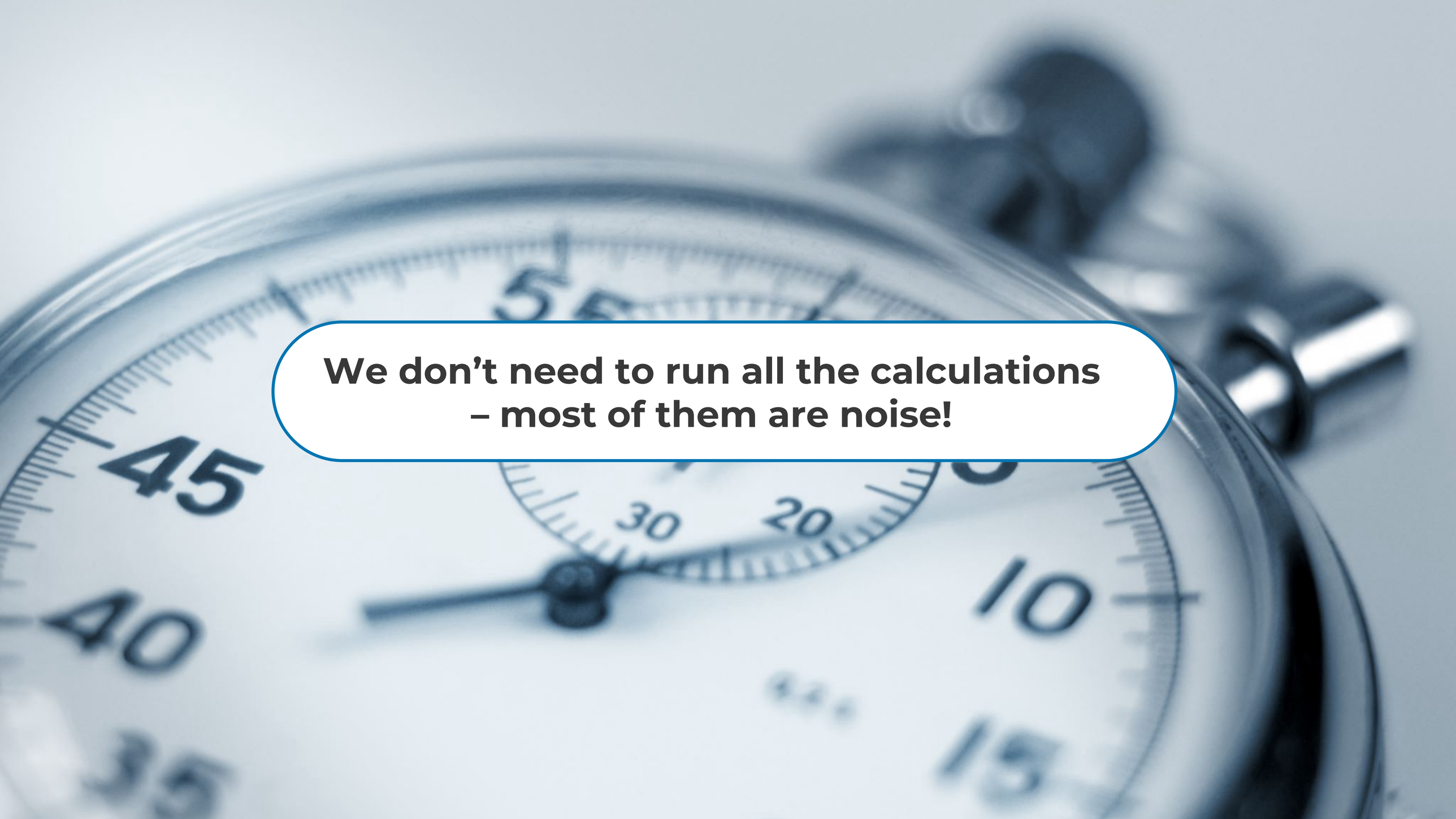


Our calculations will take hours… for each donor result

0: 2% 1: 10% 2: 7%	M 0% 01:ABGEP 01:ABGEP	A 100% 01:NRAR 35:01:01	A 100% 01:NRAR 35:01:01	L 0% 04:03 11:01	P 10% 03:02:01 03:01:01	P 11:01:01G 04:03:01G	7450 IT-Milano	Female	25	O+

The sets of genes involved differ in likelihood by 10 orders of magnitude


f(x)
Exponential
"Heavy Tailed"
x


We don’t need to run all the calculations
– most of them are noise!

Softwire

Takeaways

1. Flip the problem around
2.
3.

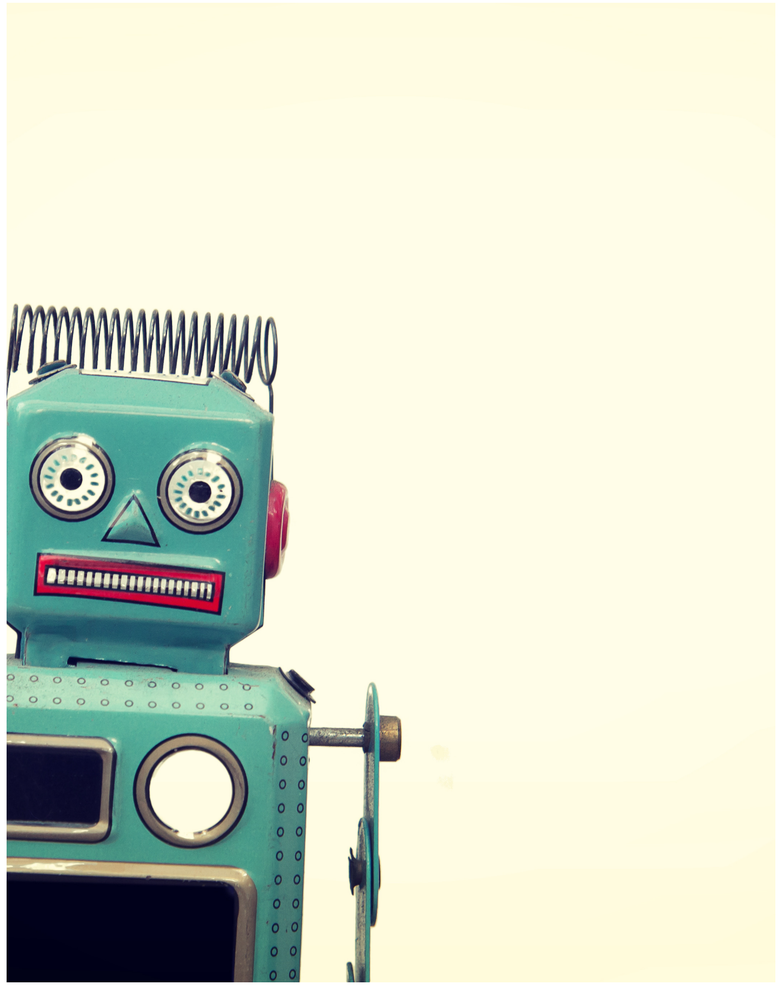
Softwire

Takeaways

1. Flip the problem around
2. Be prepared to trade precision for speed
3.

Softwire

Takeaways



1. Flip the problem around
2. Be prepared to trade precision for speed
3. The solution often isn’t a technical one

Softwire

Takeaways

1. Understand your domain
2.
3.
4.

Softwire

Takeaways

1. Understand your domain
2. Ask the right questions
3.
4.

Softwire

Takeaways

1. Understand your domain
2. Ask the right questions
3. Seek out and challenge NFRs
4.

Softwire

Takeaways

1. Understand your domain
2. Ask the right questions
3. Seek out and challenge NFRs
4. Join the register!

Ben Below

Technical Principal – Softwire

Email: benjamin.below@softwire.com

LinkedIn: www.linkedin.com/in/ben-below/

Thank you