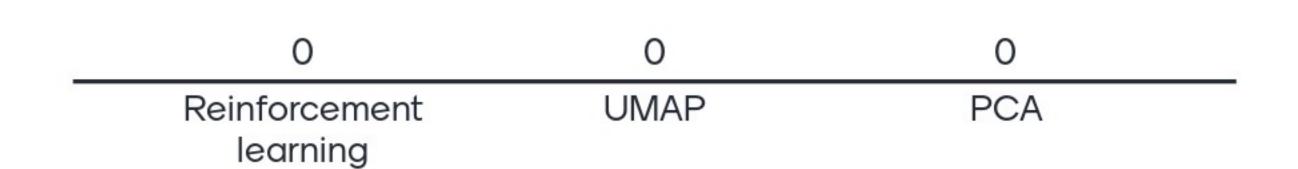


### Intro to AI/ML

Let's recap the basic concepts from m0



## I want to find groups of similar molecules in my dataset, what will I do?





# For unsupervised learning, I need labelled data

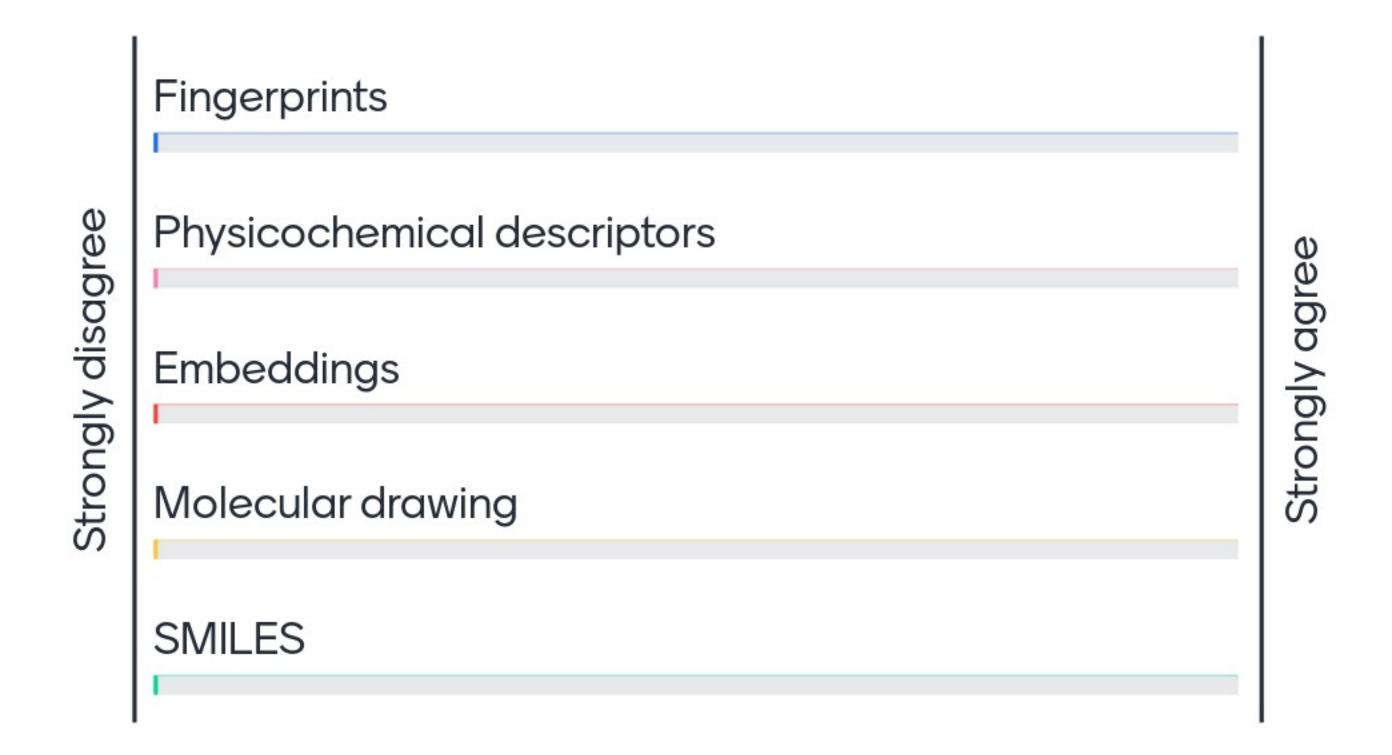




## If I have a small dataset and a very complex model, I might risk having...



# What can I use to describe molecules for ML?



# Which of the following is NOT true about SMILES?

0

It provides a way to represent chemical structures as text strings

0

It can accurately represent 3D structures

0

0

It is used in cheminformatics databases and tools Different molecules can have more than one valid SMILES representation



## Which databases can I use to find bioassay data?

# Which physicochemical properties can help me analyse a chemistry dataset?

Natural products are more diverse

Clusters of molecules are good for ML

The Lipinski rule of 5 is helpful to identify good molecules

## How can Al help in our exercise to select candidates?

#### What is the Ersilia Model Hub

<u>-</u>	Ο	0	0	
	A platform	A platform	A platform	
	to build Al	providing Al	providing	
	models	models	chemistry	
			datasets	



## What kind of Al models we find in the Ersilia Model Hub?

0	0	0	0
Models published in academic literature	Models with proprietary licenses	Models developed by Ersilia	Models using cloud servers



#### Selected Models from the Ersilia Model Hub!

#### Antimalarial models available

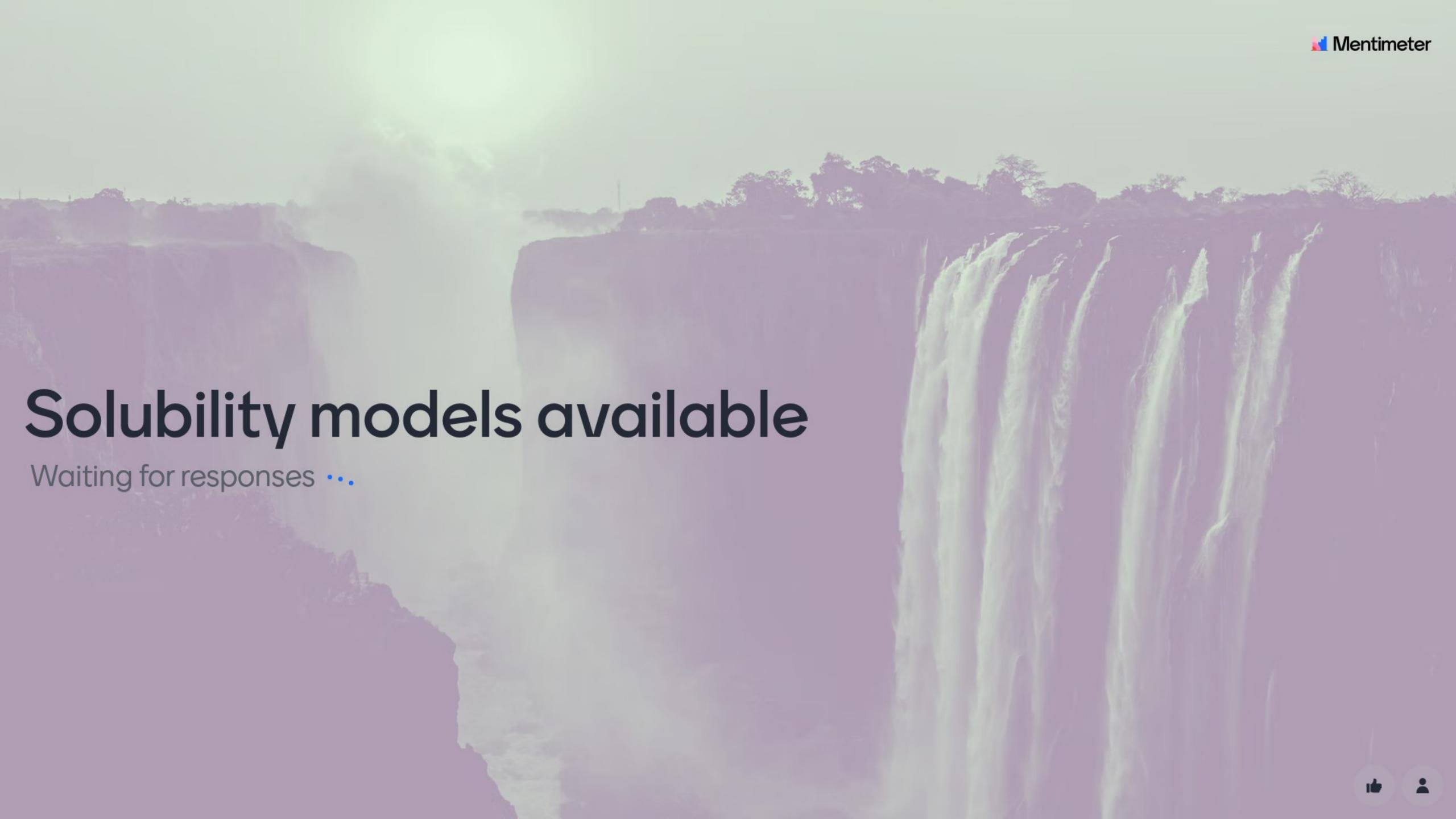


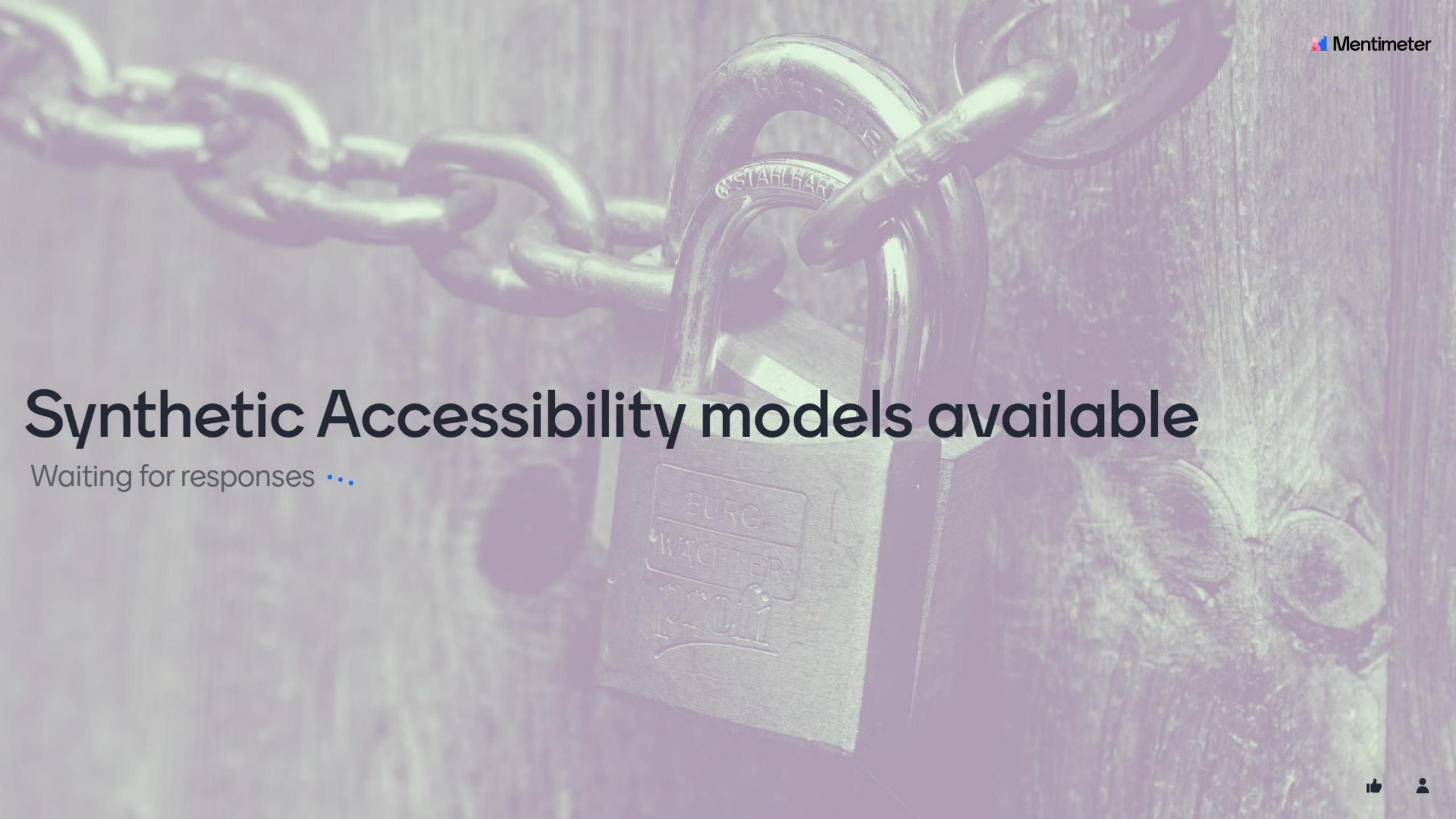


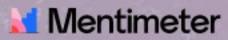


### Cardiotoxicity models available









### Natural product scoring models available