Towards closed-loop automation of pharmaceutical formulations

RSC Formulation 4.0 Conference 2024

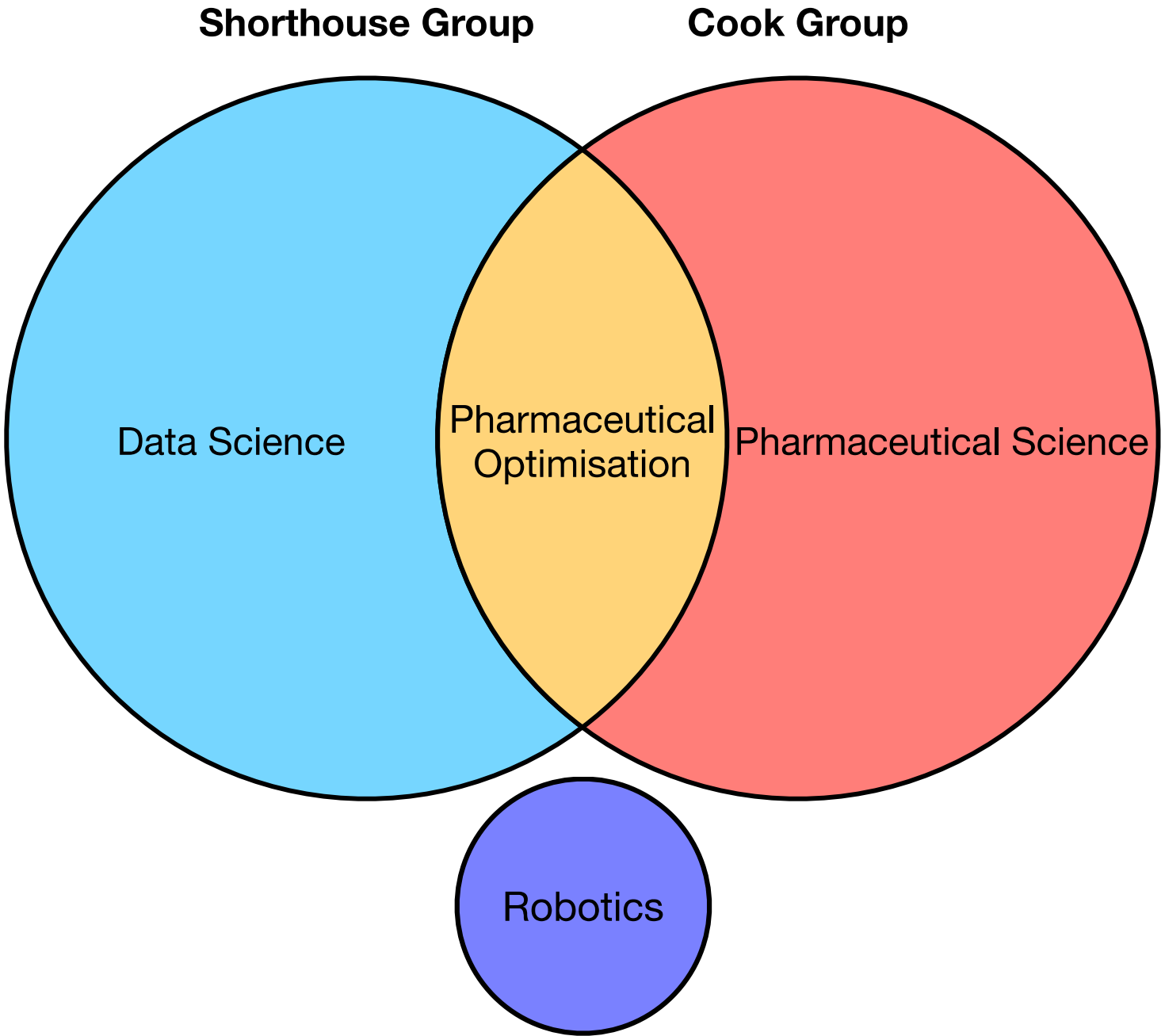
Dr David Shorthouse: d.shorthouse@ucl.ac.uk

Dr Michael Cook: m.cook@ucl.ac.uk

UCL School of Pharmacy

Project launch: “Artificial intelligence coupled to automation for accelerated medicine design”

Engineering and Physical Sciences Research Council


Shorthouse Group
Cook Group
Data Science
Pharmaceutical Optimisation
Pharmaceutical Science
Robotics

Traditional formulation

- Time consuming
- Datasets often unavailable
- Inconsistent quality
- Reliant on dwindling expertise and skilled operators (GAP0018 – SIP, Industrial Strategy 2017)
- Design space inherently constrained by human intuition and quantities
- TPP/mTPP/qTPP – multiple endpoints.
- Quality reporting requires expertise

A revolution in the chemical sciences?

CHEMISTRY WORLD

Your new labmate does almost 700 reactions in eight days – and it's a robot

Google AI and robots join forces to build new materials

Tool from Google DeepMind predicts nearly 400,000 stable substances, and an autonomous system learns to make them in the lab.

AstraZeneca iLab: The automated lab of the future

Autonomous polymer synthesis delivered by multi-objective closed-loop optimisation†

Stephen T. Knox‡, Sam J. Parkinson‡§, Clarissa Y. P. Wilding, Richard A. Bourne and Nicholas J. Warren*

School of Chemical and Process Engineering, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, West Yorkshire, UK. E-mail: n.warren@leeds.ac.uk

An autonomous portable platform for universal chemical synthesis

J. Sebastián Manzano, Wenduan Hou, Sergey S. Zalesskiy, Przemyslaw Frei, Hsin Wang, Philip J. Kitson & Leroy Cronin

Nature Chemistry **14**, 1311–1318 (2022) | Cite this article

Operator-independent high-throughput polymerization screening based on automated inline NMR and online SEC†

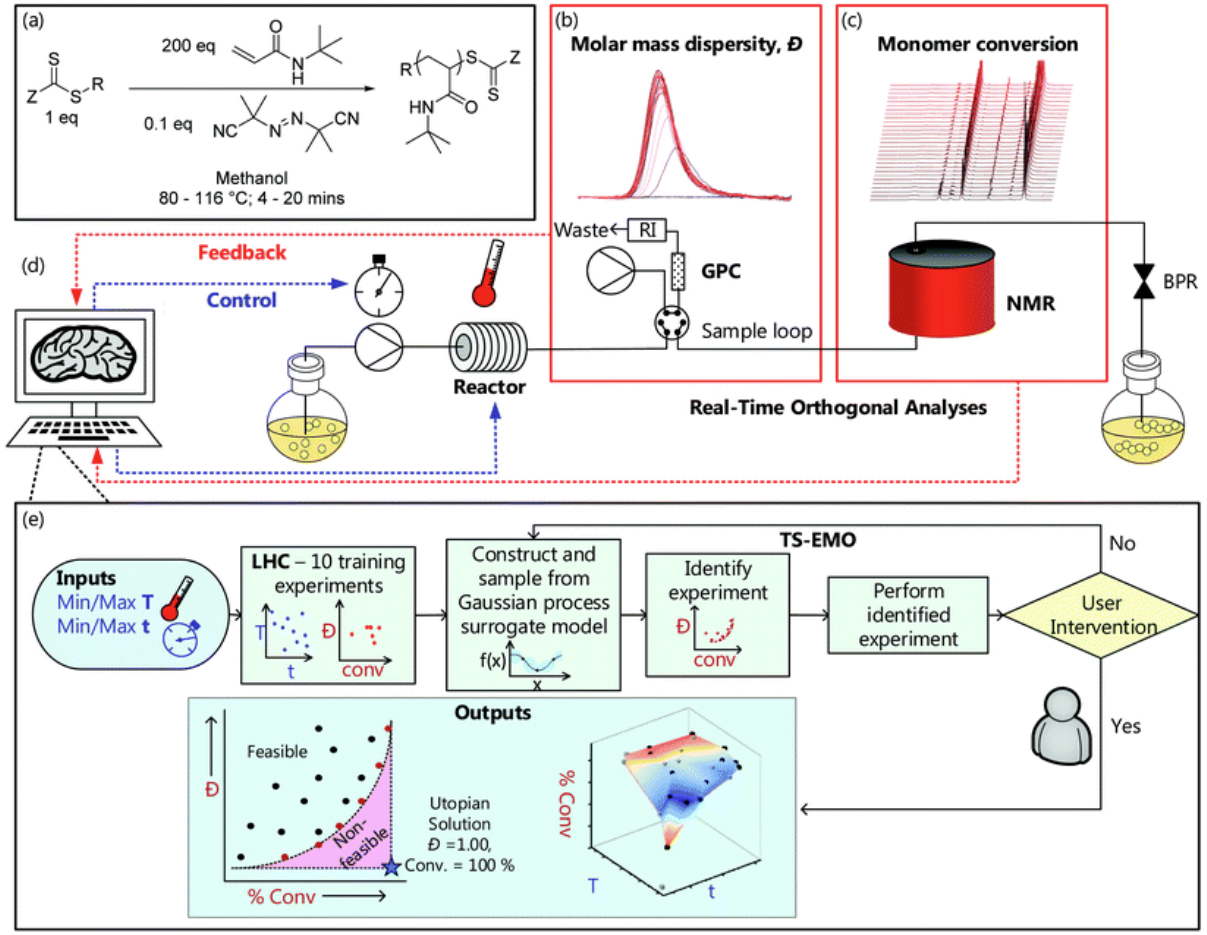
Joren Van Herck, Iyomali Abeysekera, Axel-Laurenz Buckinx, Kewei Cai, Jordan Hooker, Kirti Thakur, Emma Van de Reydt, Pieter-Jan Voorter, Dries Wyers and Tanja Junkers*

Polymer Reaction Design Group, School of Chemistry, Monash University, 19 Rainforest Walk, Clayton, Victoria 3800, Australia. E-mail:

A revolution in the chemical sciences?

Burger et al, Nature, 2020, 583, 237–241

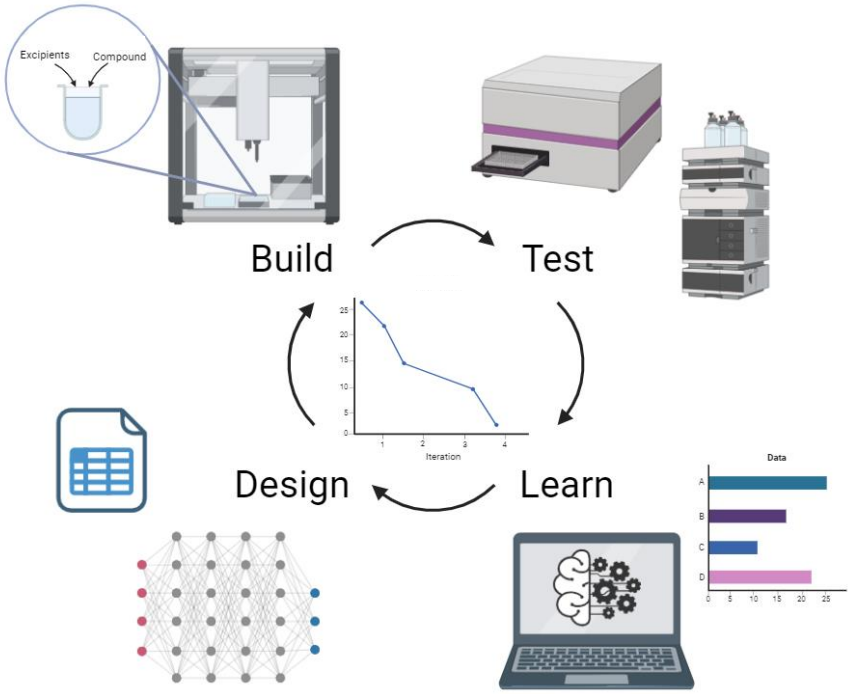
A revolution in the chemical sciences?



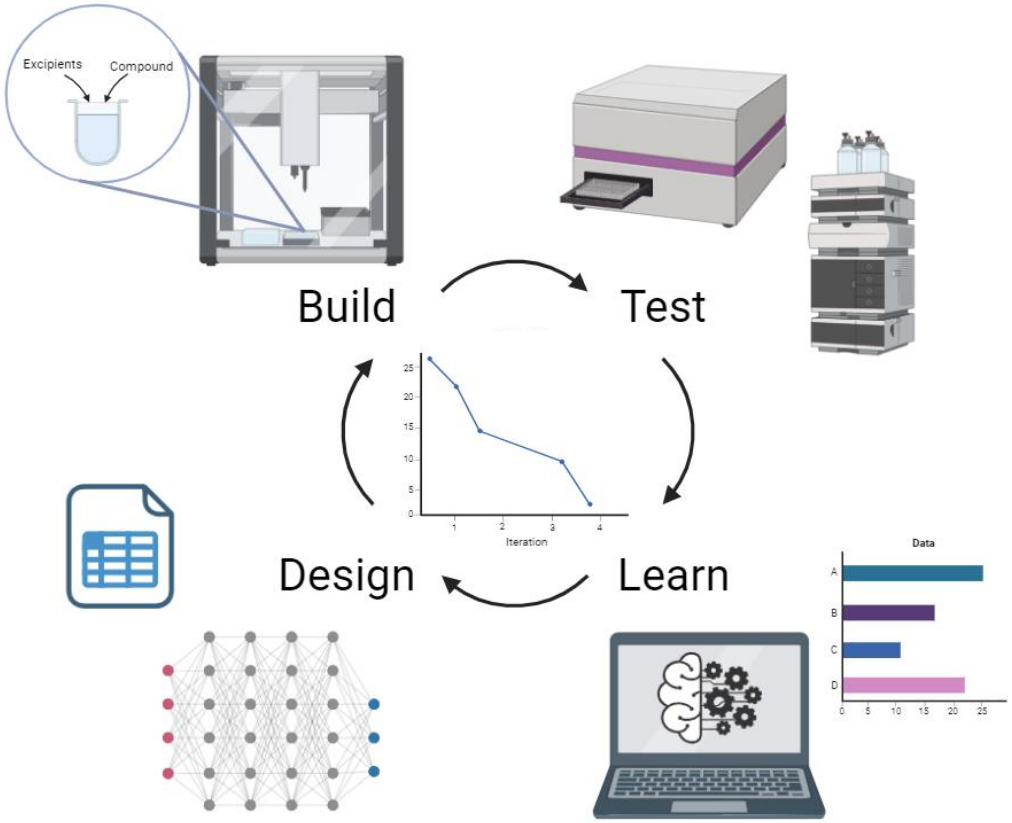
Knox et al, Polym. Chem., 2022, 13, 1576-1585

A vision for modern formulation

- Bespoke automated pipelines for a particular dosage form with closed-loop capability
- **Full** exploration of excipient space
- Multiple endpoints matched to TPP
- Standardised reporting with data integrity



Where are we?


Excipients
Compound
Build
Test
Design
Learn
Iteration
Data
A
B
C
D

Combinatorial explosions

Combining things in different concentrations gets very complex very quickly.

If we have **2** excipients that can be combined in **11** different concentrations (0 to 10%) then we have:

11 x 11 = **121** different ways of combining them.

This increases rapidly – we multiply this by 11 each time we add a new excipient.

If we have **10** excipients that can be combined in **11** different concentrations (0 to 10%) then we have:

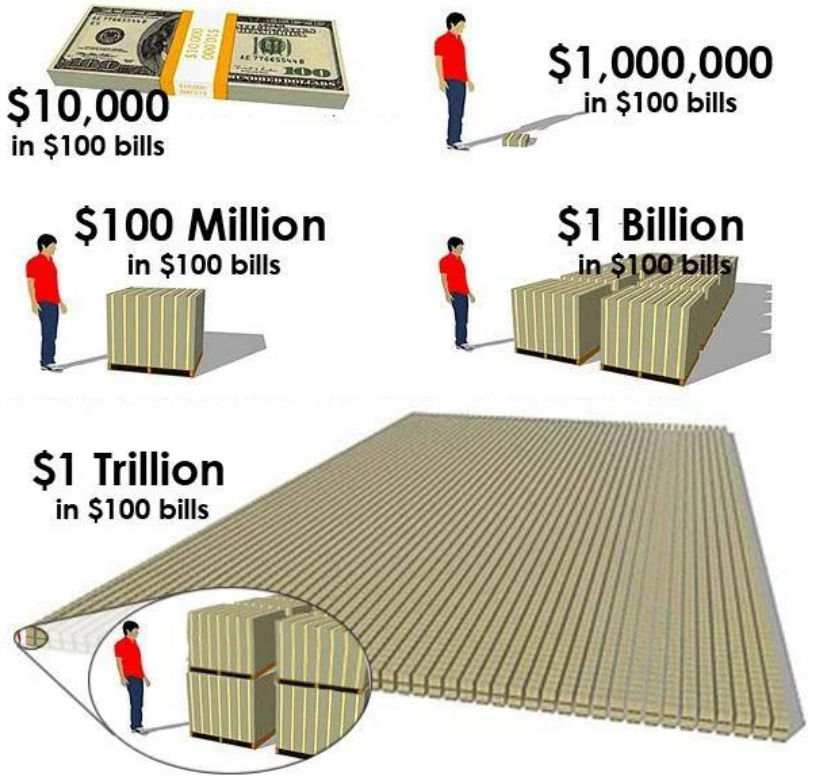
11 x 11 x 11 x 11 x 11 x 11 x 11 x 11 x 11 x 11 or 11^{10} = **25,937,424,601** ways to combine them.

Combinatorial explosions

One Million seconds is ~12 days

One Billion seconds is ~32 years

One trillion seconds is ~30,000 years



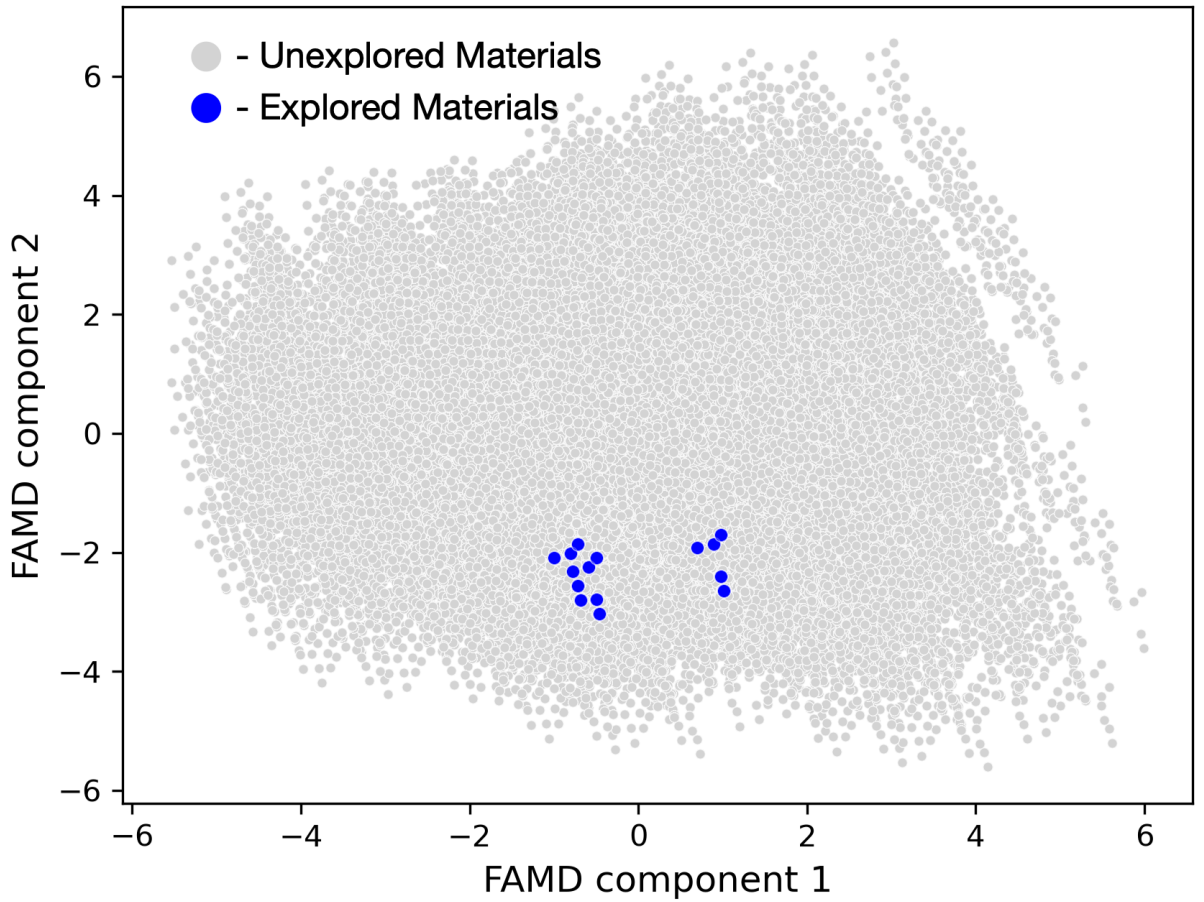
Our current exploration of formulation space is limited

Mucoadhesive polymer formulations for drug delivery:

- 101 formulations published.

- Assuming all possible combinations of every parameter in the dataset (unrealistic), there are ~100 billion possible polymer formulations.

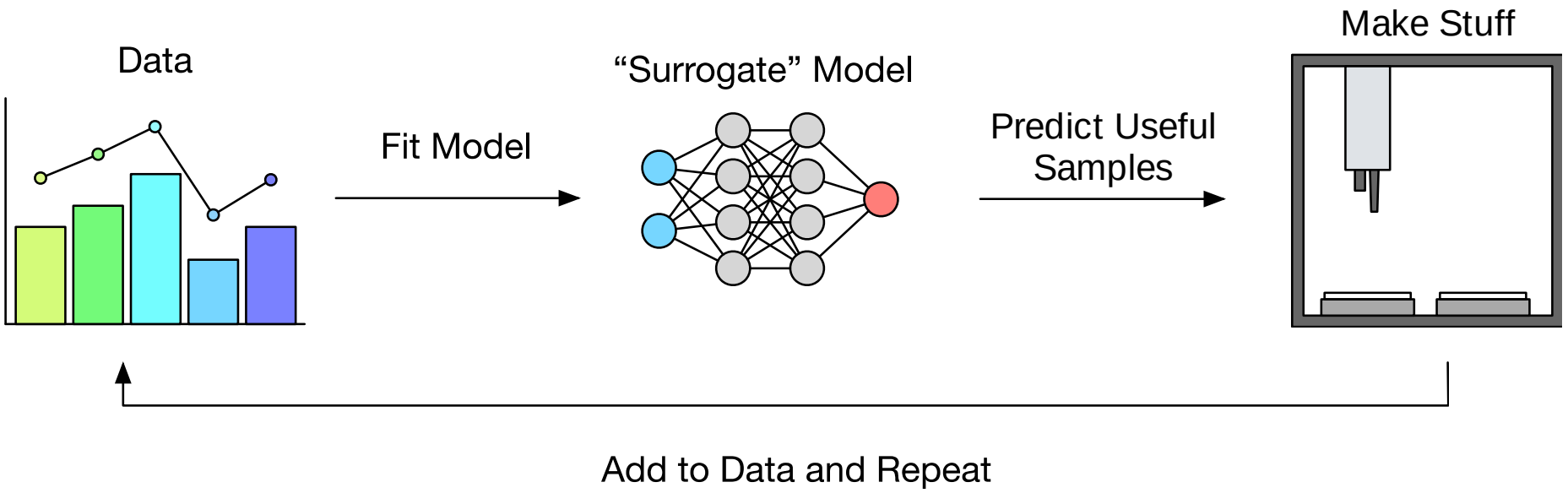
- 0.0009% of the state space has been explored.



Cook, Shorthouse. Under review

Bayesian Optimisation

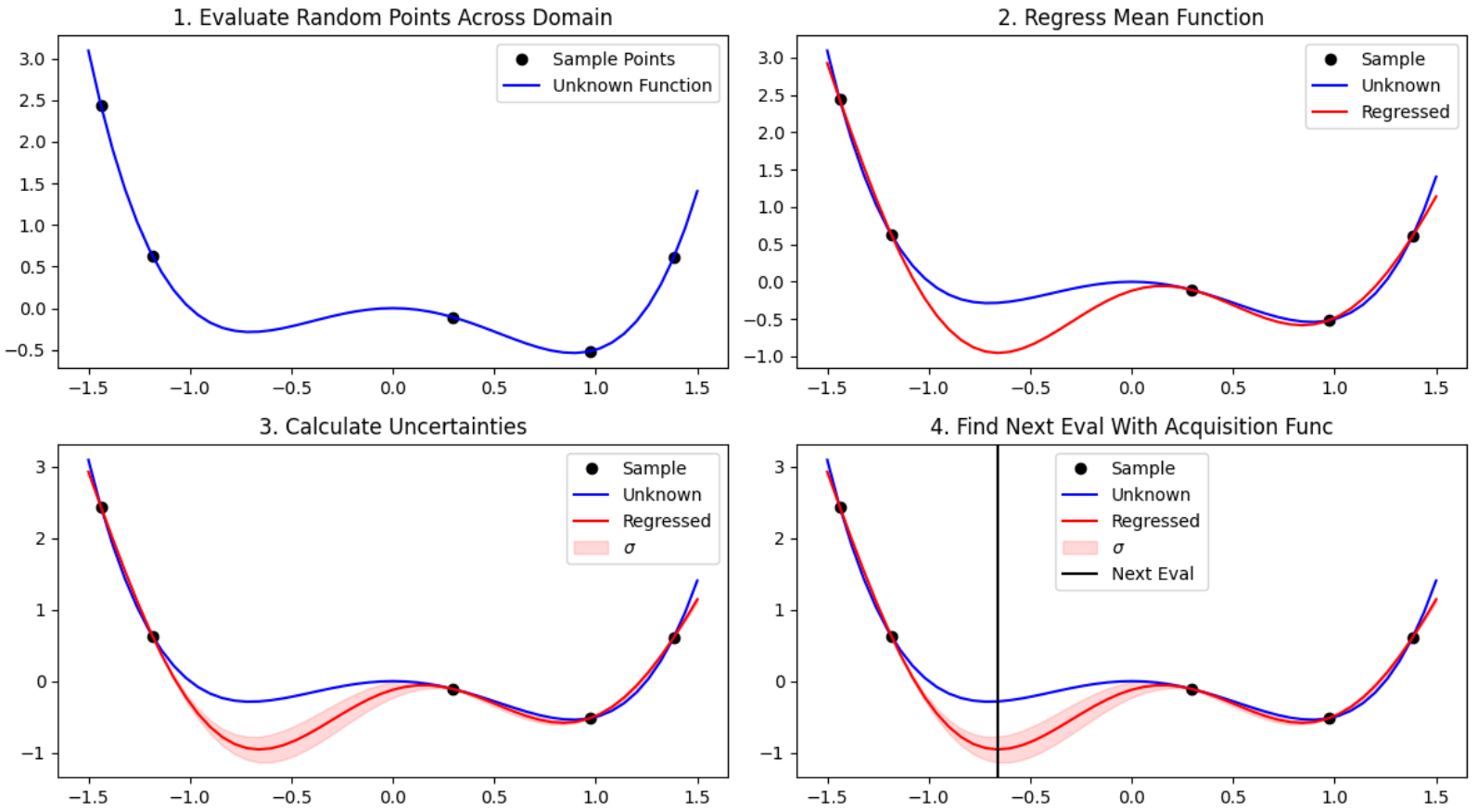
Bayesian Optimisation offers a method for exploring these systems in an intelligent and efficient way.



Bayesian Optimisation

The surrogate model generally relies on the **uncertainty of its predictions** to recommend the next points for study.

There are multiple mathematical methods available for calculating the “best” next data to collect.

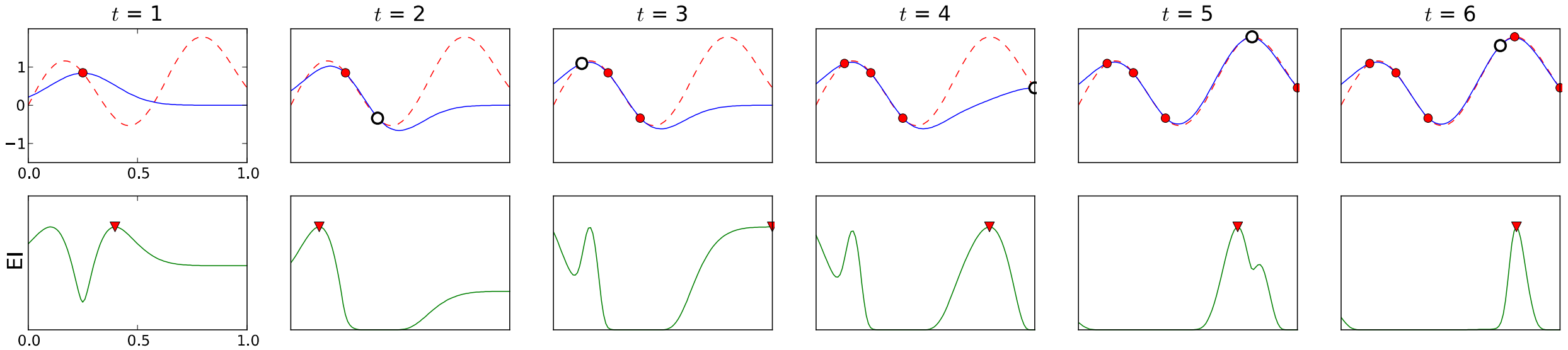


Medium.com

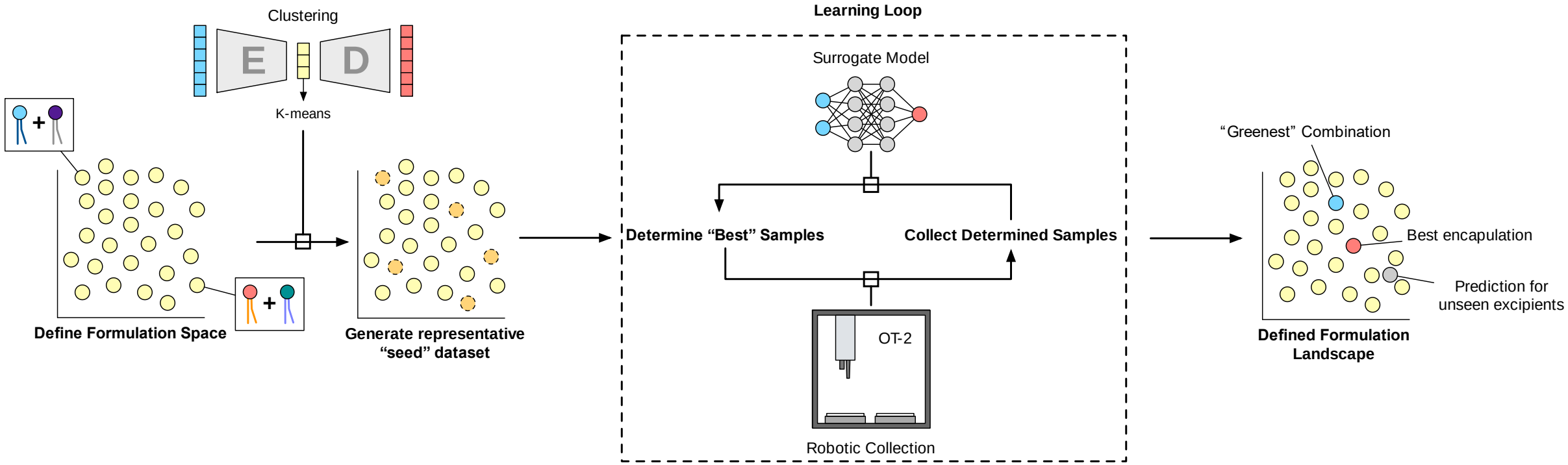
Bayesian Optimisation

The surrogate model generally relies on the **uncertainty of its predictions** to recommend the next points for study.

There are multiple mathematical methods available for calculating the "best" next data to collect.



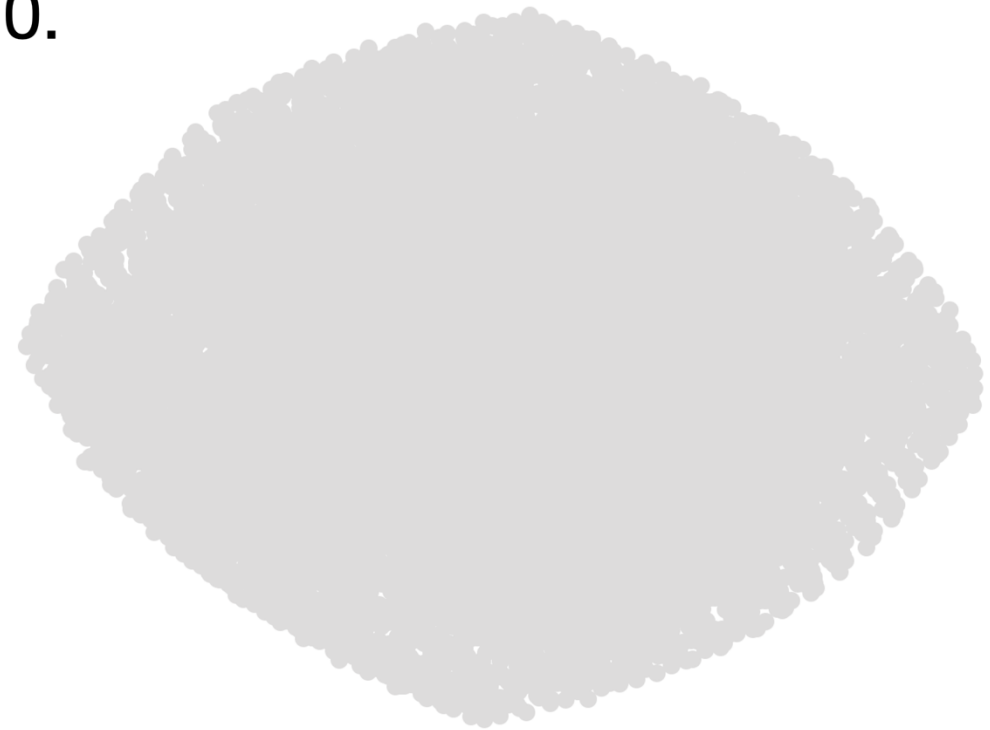
Brochu et al. Arxiv (2010). https://doi.org/10.48550/arXiv.1012.2599

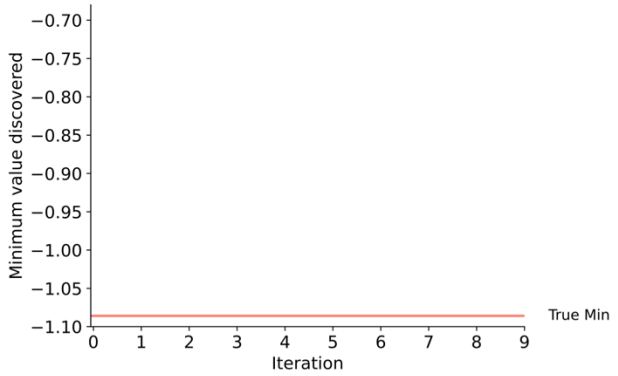

Clustering
E
D
K-means
Define Formulation Space
Generate representative "seed" dataset
Learning Loop
Surrogate Model
Determine "Best" Samples
Collect Determined Samples
OT-2
Robotic Collection
"Greenest" Combination
Best encapulation
Prediction for unseen excipients
Defined Formulation Landscape

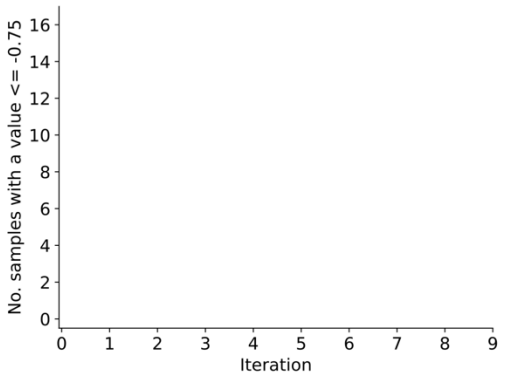
An example:

- System of 6 excipients with 11 potential concentrations each.
- **17 million** possible combinations.
- We “seed” our system with 96 samples (3 x 96 well plates).
- We then run batches of 32 (one plate with 3 repeats) samples.
- We get near to a theoretic optimum within 400/17,000,000 (**0.002%** of) samples .

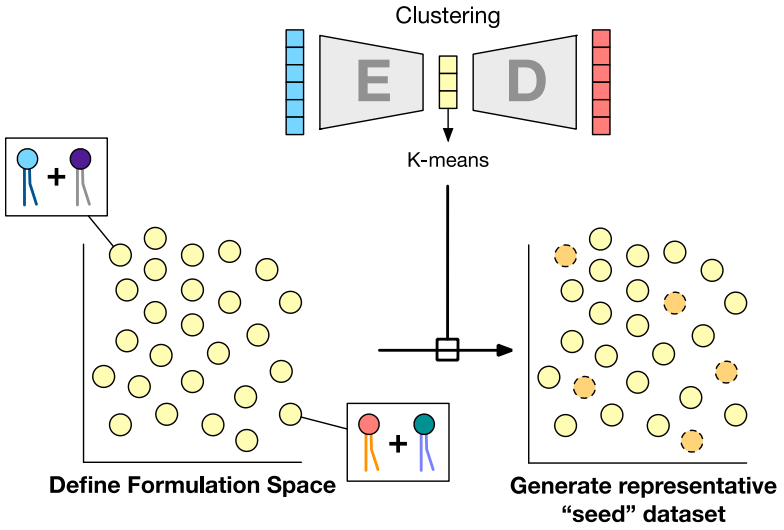
0.



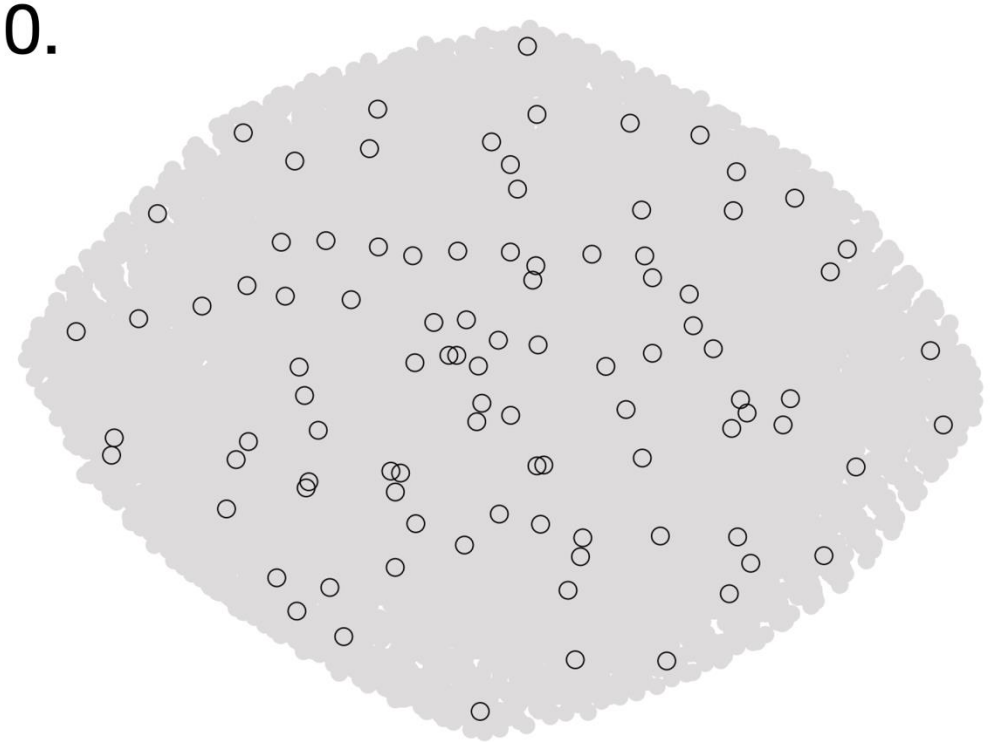


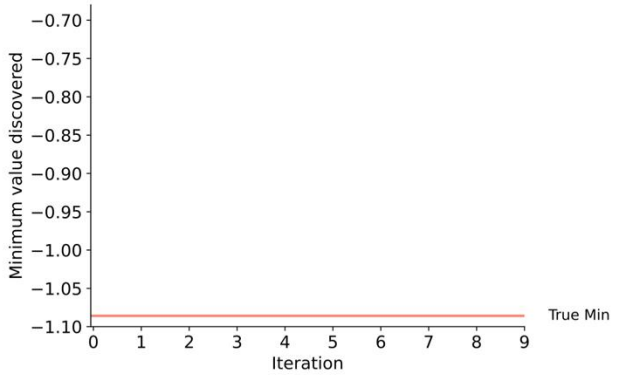


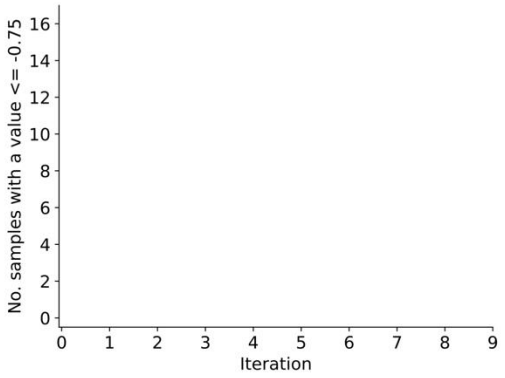
An example:


Clustering
E
D
K-means
Define Formulation Space
Generate representative
“seed” dataset

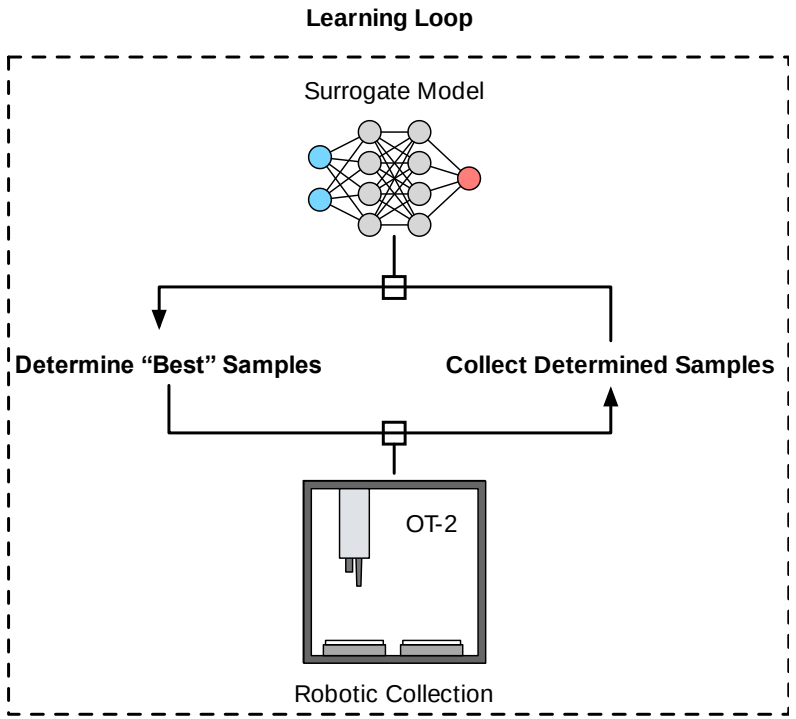
0.



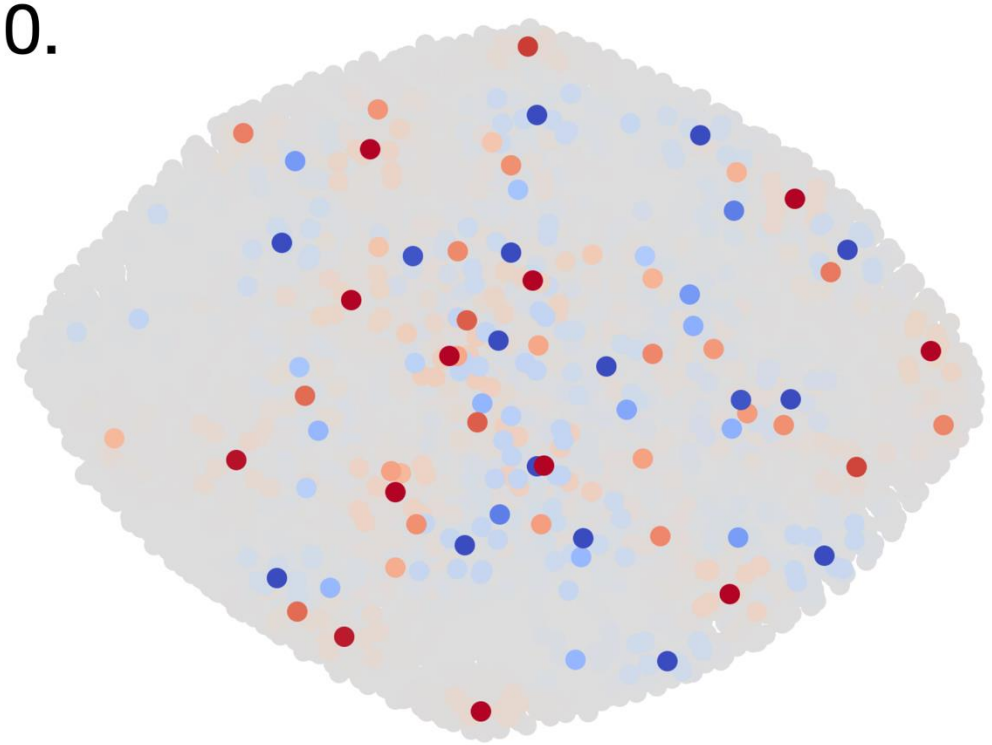

Minimum value discovered
−0.70
−0.75
−0.80
−0.85
−0.90
−0.95
−1.00
−1.05
−1.10
0 1 2 3 4 5 6 7 8 9
Iteration
True Min

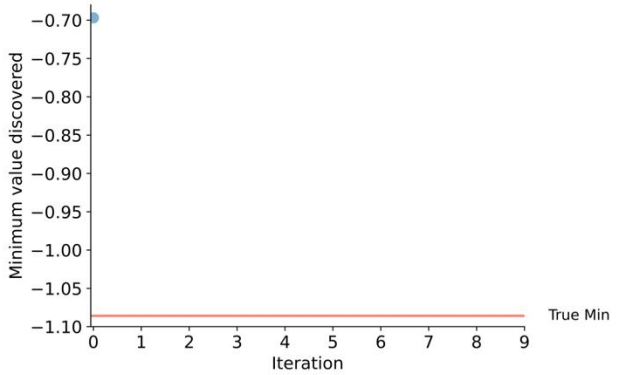

No. samples with a value <= -0.75
16
14
12
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9
Iteration

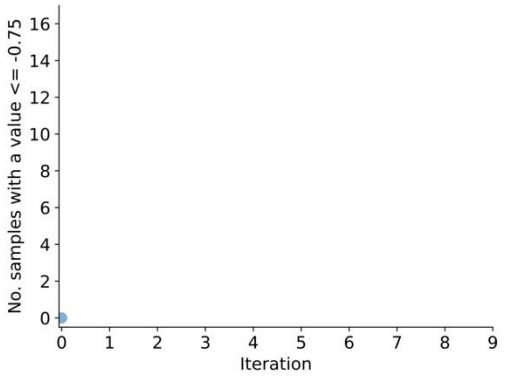
An example:


Learning Loop
Surrogate Model
Determine “Best” Samples
Collect Determined Samples
OT-2
Robotic Collection

0.

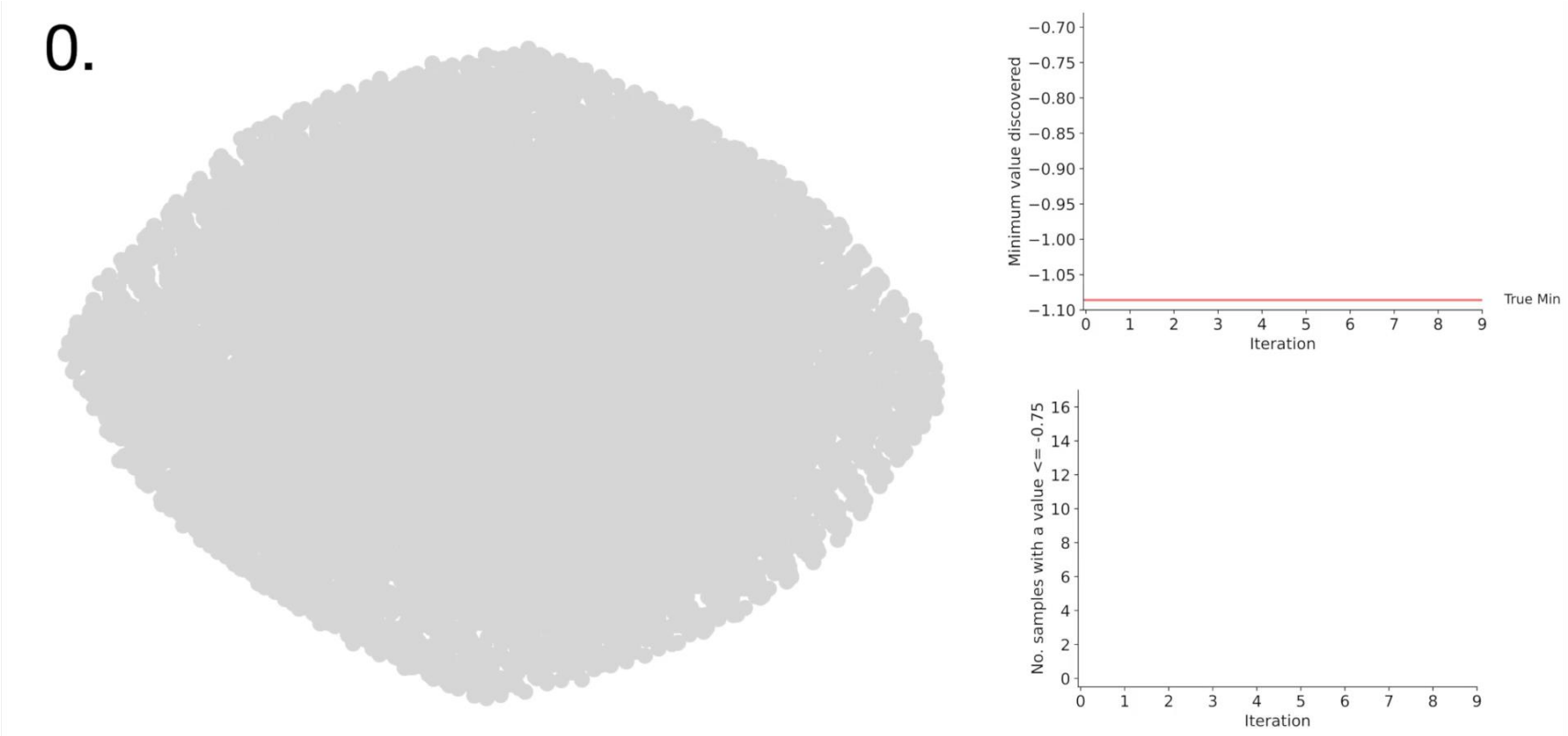



Minimum value discovered
−0.70
−0.75
−0.80
−0.85
−0.90
−0.95
−1.00
−1.05
−1.10
0 1 2 3 4 5 6 7 8 9
Iteration
True Min


No. samples with a value <= -0.75
16
14
12
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9
Iteration

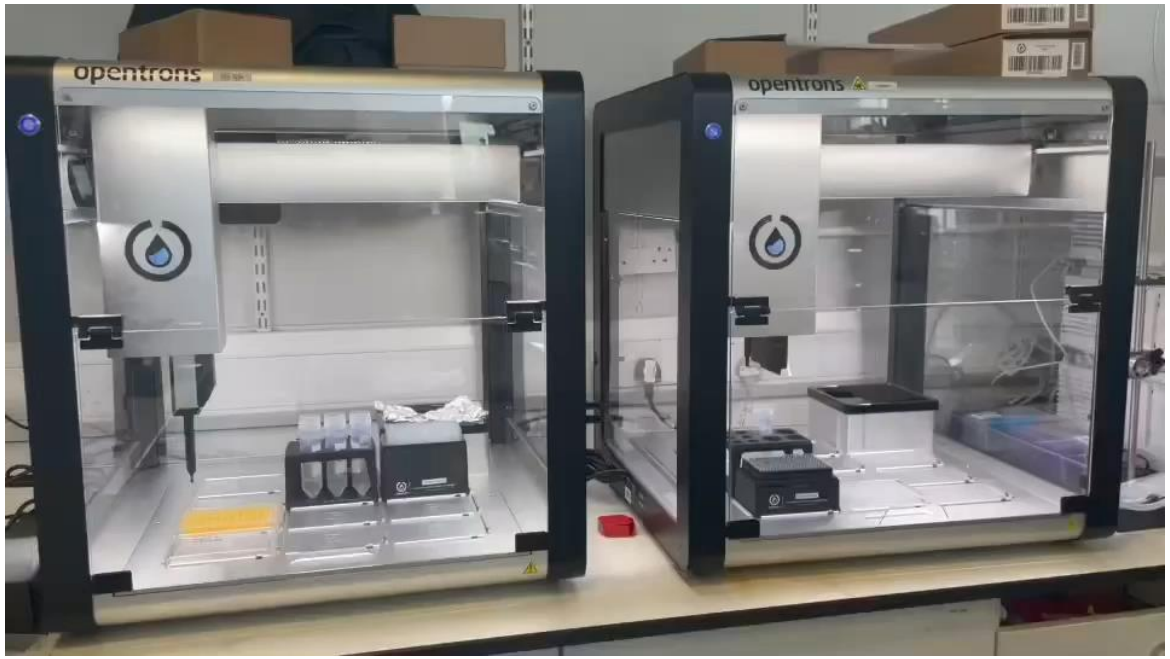
An example:

- System of 6 excipients with 11 potential concentrations each
- **17 million** possible combinations
- We “seed” our system with 96 samples (3 x 96 well plates)
- We then run batches of 32 (one plate with 3 repeats) samples
- We get near to a theoretic optimum within 400/17,000,000 (**0.002%** of) samples

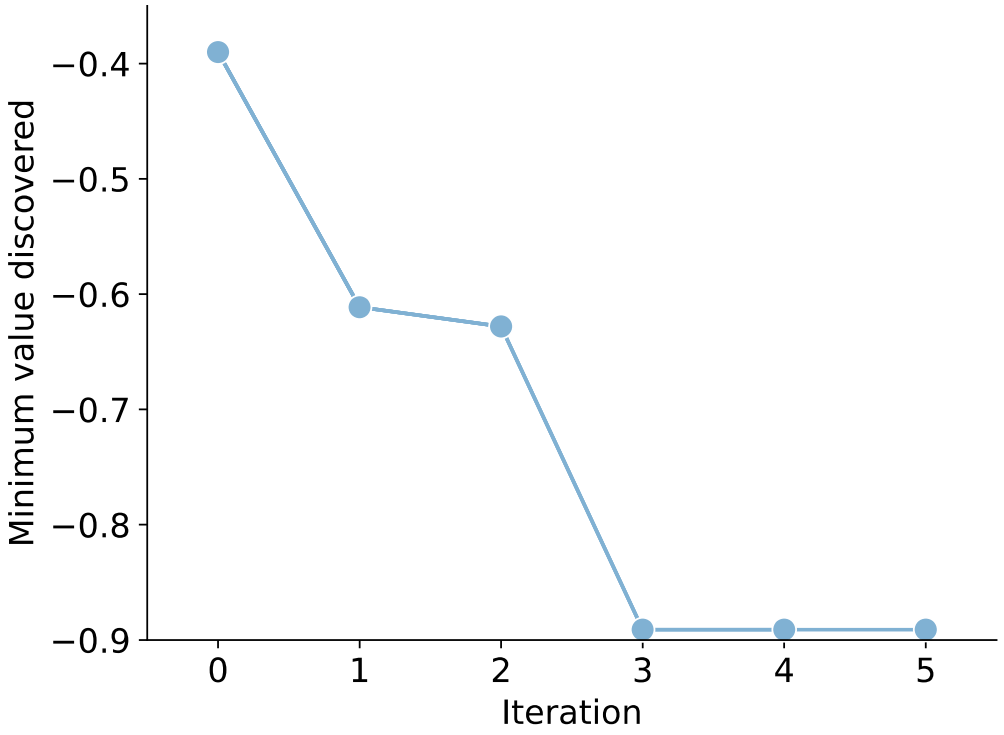


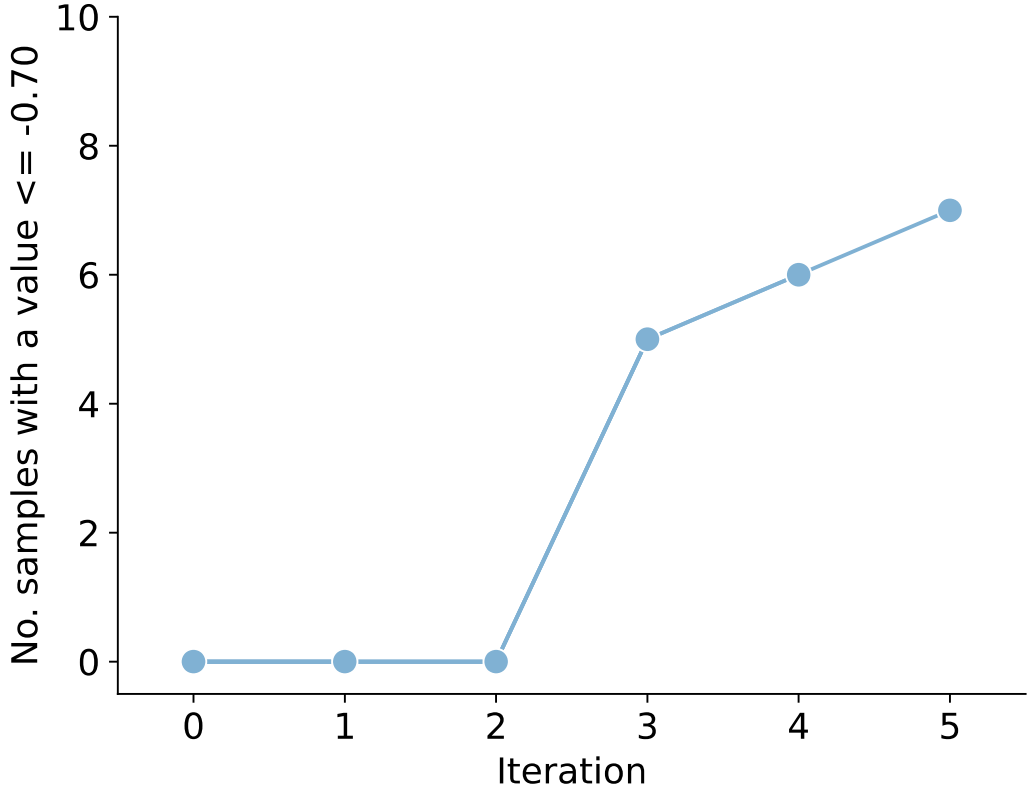
A real example:

- A target active.
- 6 possible excipients.
- 11 levels for each excipient (0 and 1-10% solutions).
- We follow the identical process as in our “toy” example system:
  - 3 x 96 well plates to seed the system.
  - 1 x 96 well plate per loop.
- We use an opentrons OT-2 to generate the samples.
- Each loop takes ~6 hours (with manual steps required).
- We measure a target endpoint with a plate reader.



A real example:





- We reach a “good” minimum value within 5 loops

- 7 different and novel combinations discovered above our threshold

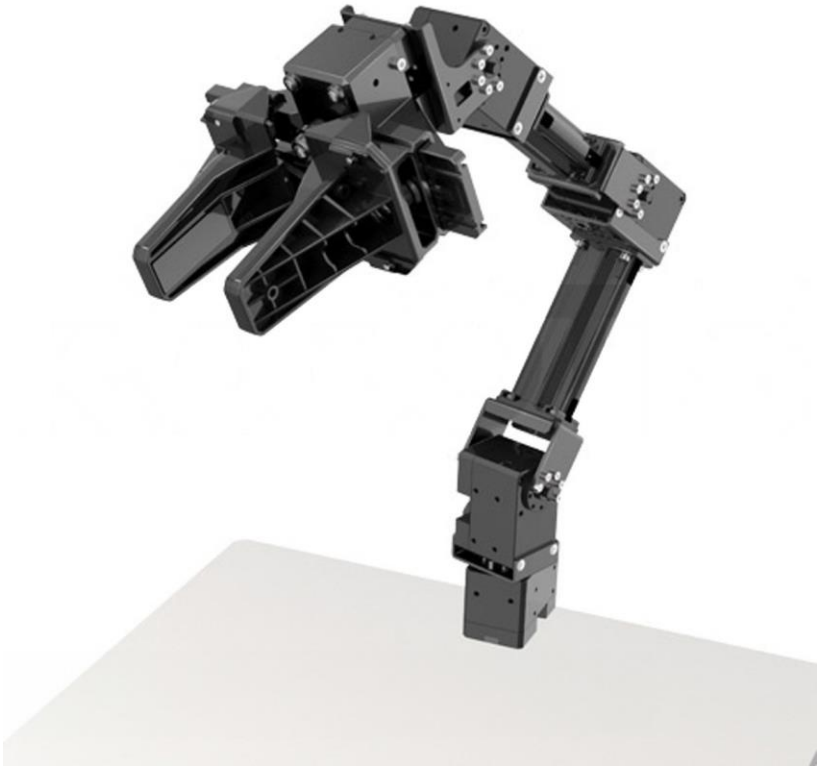
A real example:

	Surfactants				
Samples	3	5	5	5	4
	5	0	1	4	2
	5	5	3	5	5
	5	4	3	5	4
	5	4	5	1	2
	3	4	5	1	5
	4	2	2	3	1

These samples are diverse

Next Stages:

- “Close the loop”
- Industry standard QC/Assessment
- Explore efficiency of different algorithms
- Automatic reporting/collating of formulation contents
- Optimisation of multiple endpoints



Shorthouse & Cook Groups:

Antonia Gucic

Helena Ros

Shorthouse Group:

Youssef Abdalla

Sara Jamshidi Parvar

Leo Gornovskiy

Mark Kudady

Gedion Gurmahun

Rama Hassoun

Cook Group:

Eleanor Hilton

Niamh Haslett

Shuting Li

Hessam Rasooli Nia

Toyosi Akande

Jamie Summers

Marissa Taub

Li Qian

Elinda Zeqiri

Funders & Supporters:

THE ROYAL SOCIETY

UKRI Engineering and Physical Sciences Research Council

QbDVision

BAYER