UNIVERSITY OF CAMBRIDGE

# The modern day blacksmith

Gareth Conduit

Theory of Condensed Matter group

# Neural network algorithm to

Train from sparse datasets

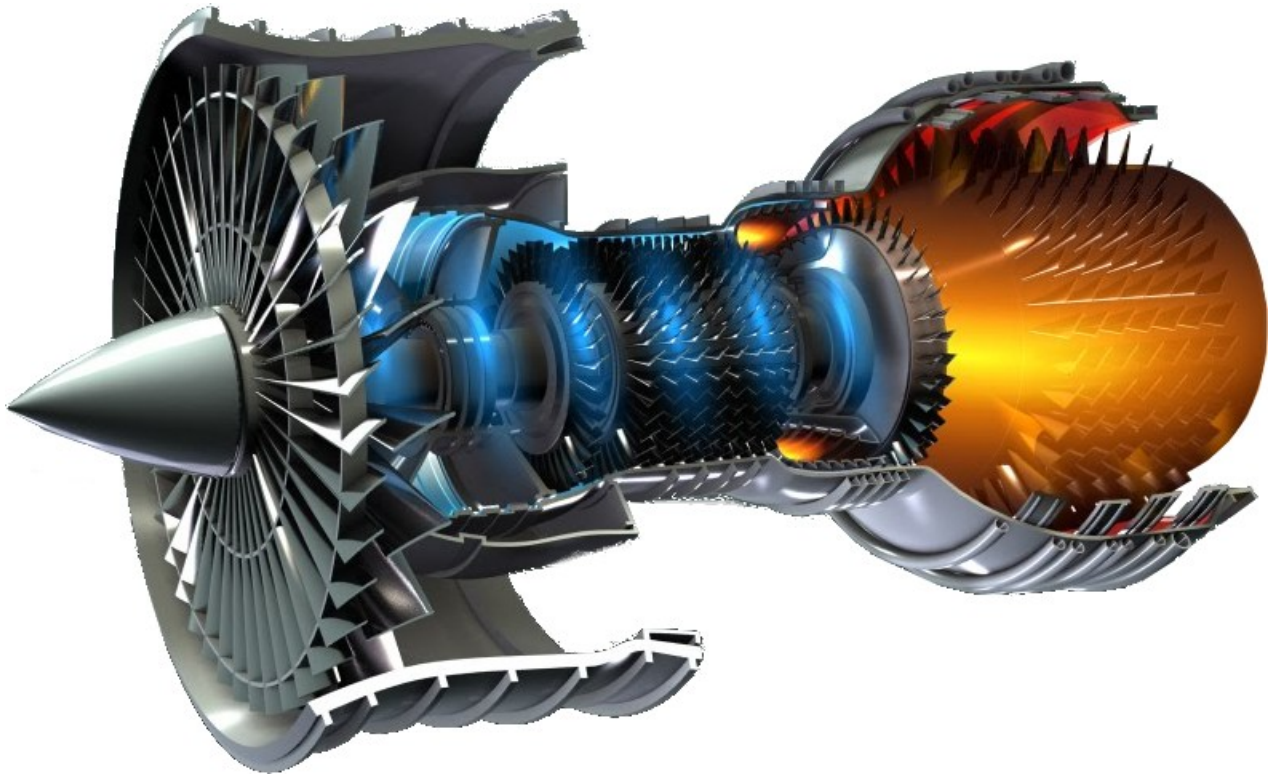
Merge simulations, physical laws, and experimental data

Reduce the need for expensive experimental development

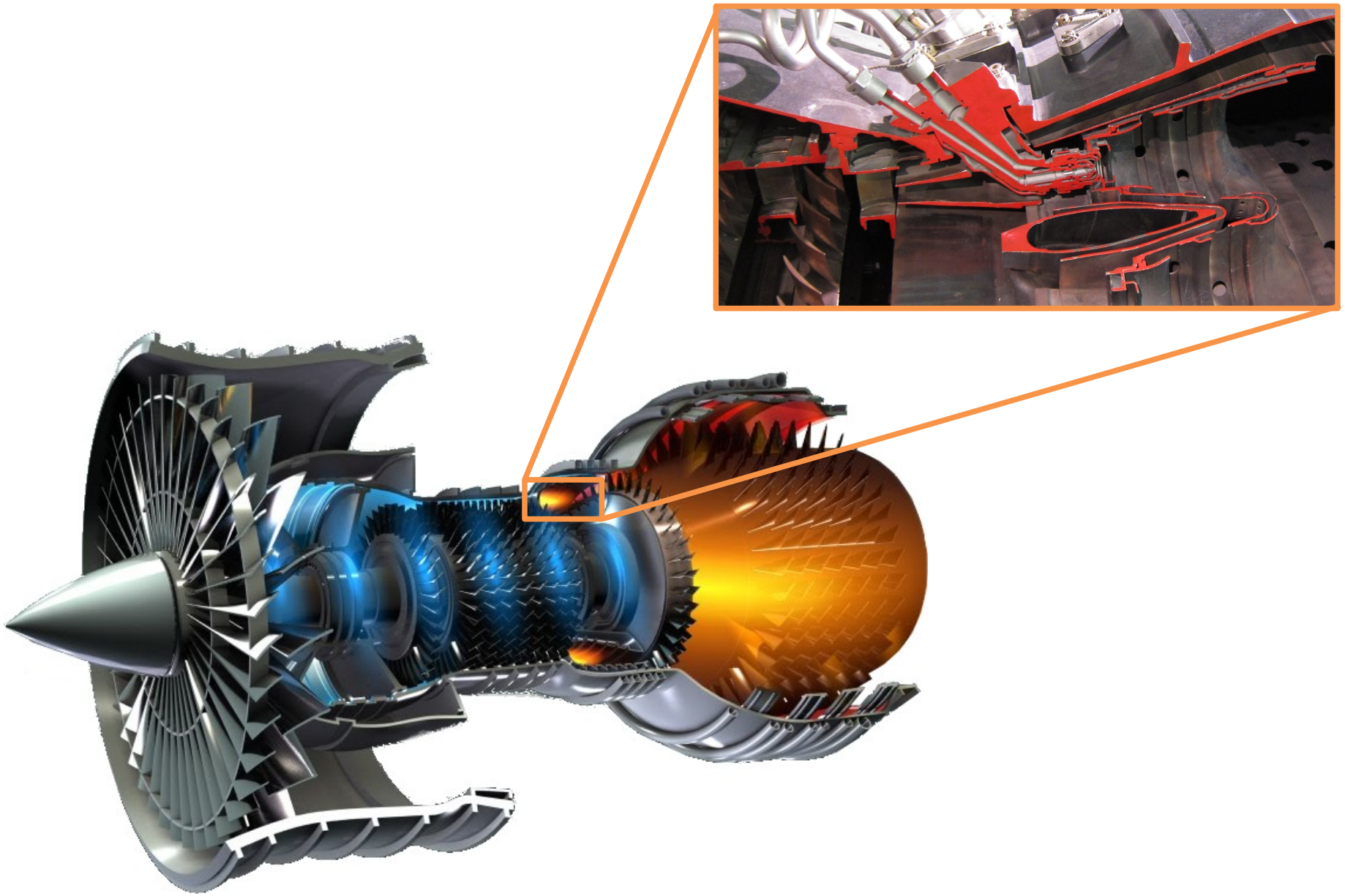
Accelerate materials and drugs discovery

Generic with proven applications in materials discovery and drug design

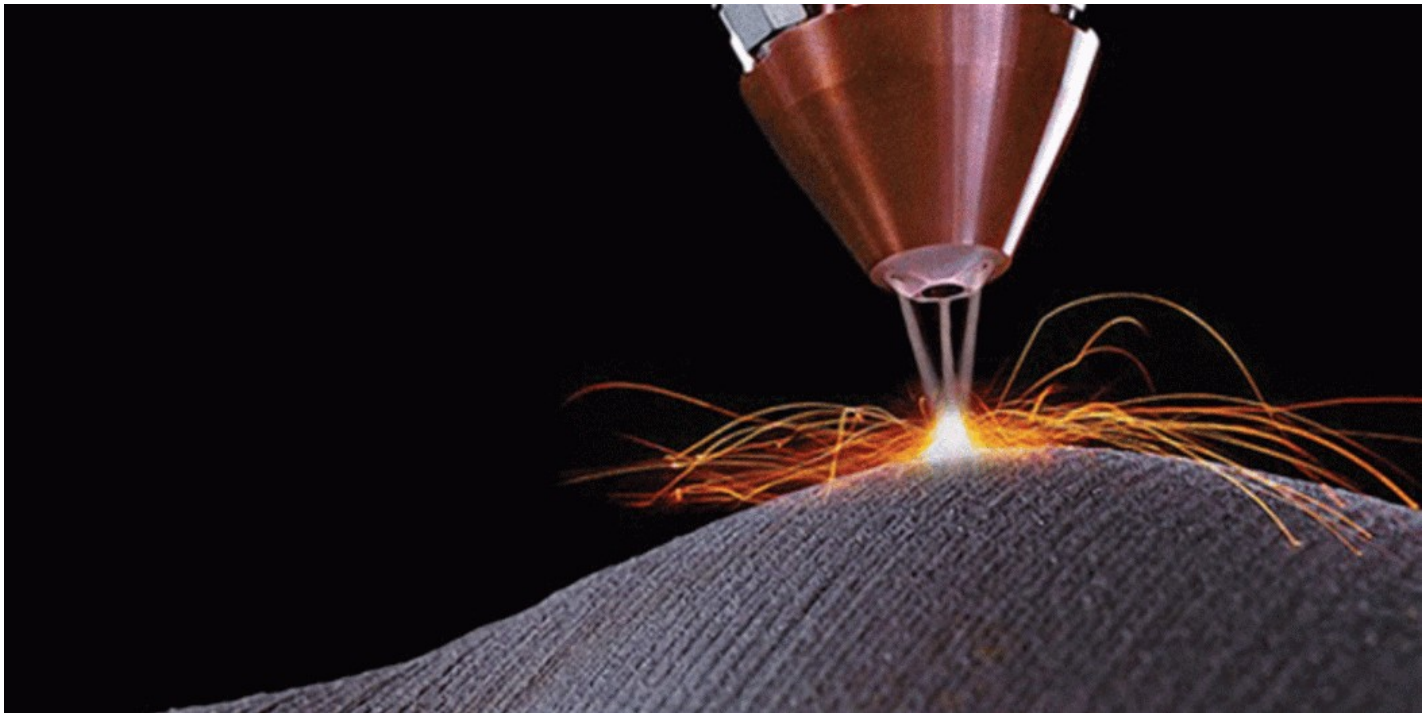
# Schematic of a jet engine



# Combustor in a jet engine



# Direct laser deposition requires new alloys

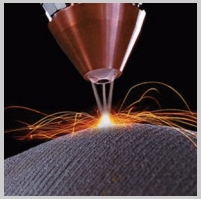


# Neural networks for materials design

## Composition

## Properties

Process



Fatigue

Welding

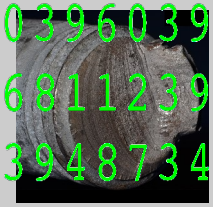
# Neural networks for materials design

**Composition**

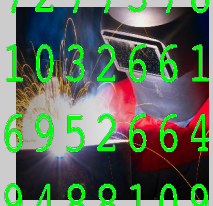
**Properties**

Process

Fatigue

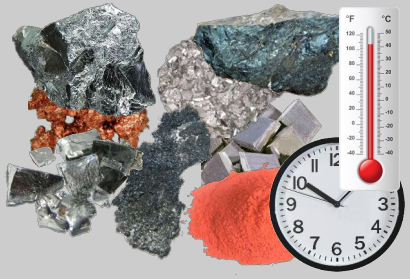


Welding



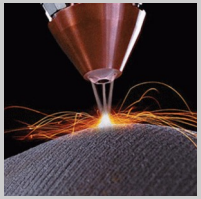
# Neural networks for materials design

## Composition



## Properties

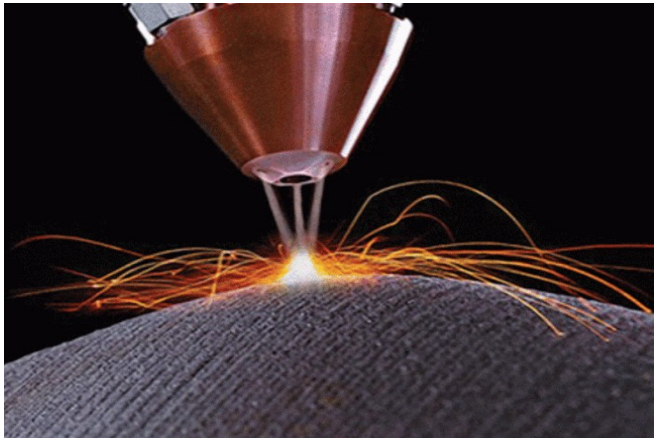
Process



Fatigue

Welding

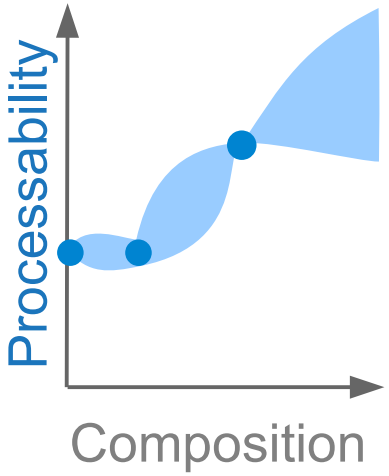
# Neural networks for materials design



Laser

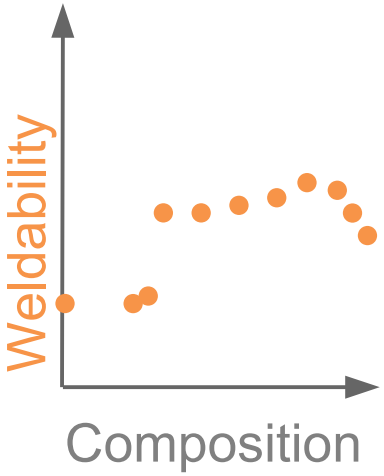
Electricity

# Insufficient data for processability

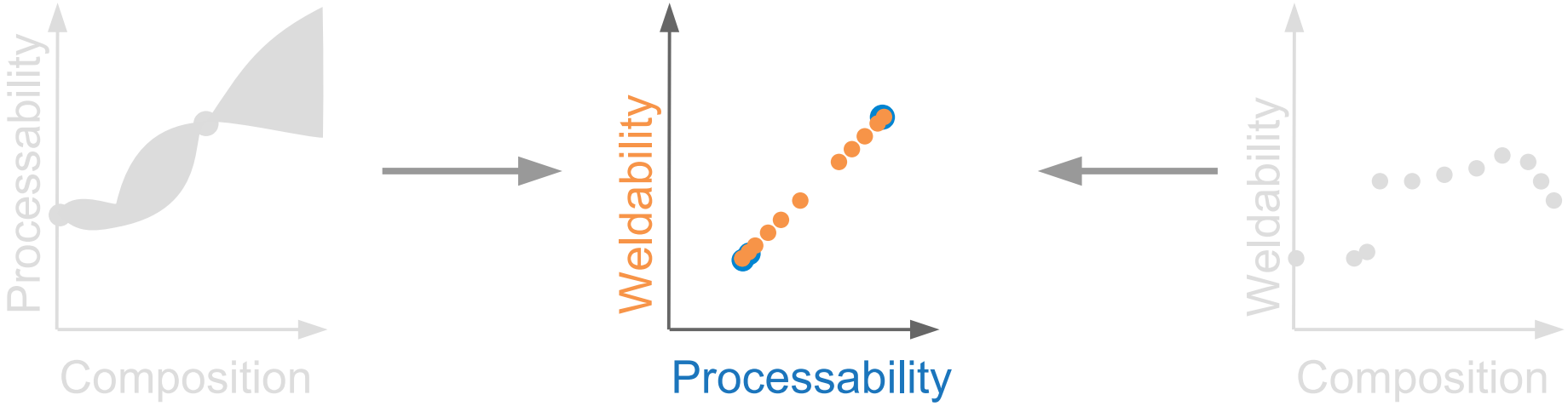


# Welding is analogous to direct laser deposition

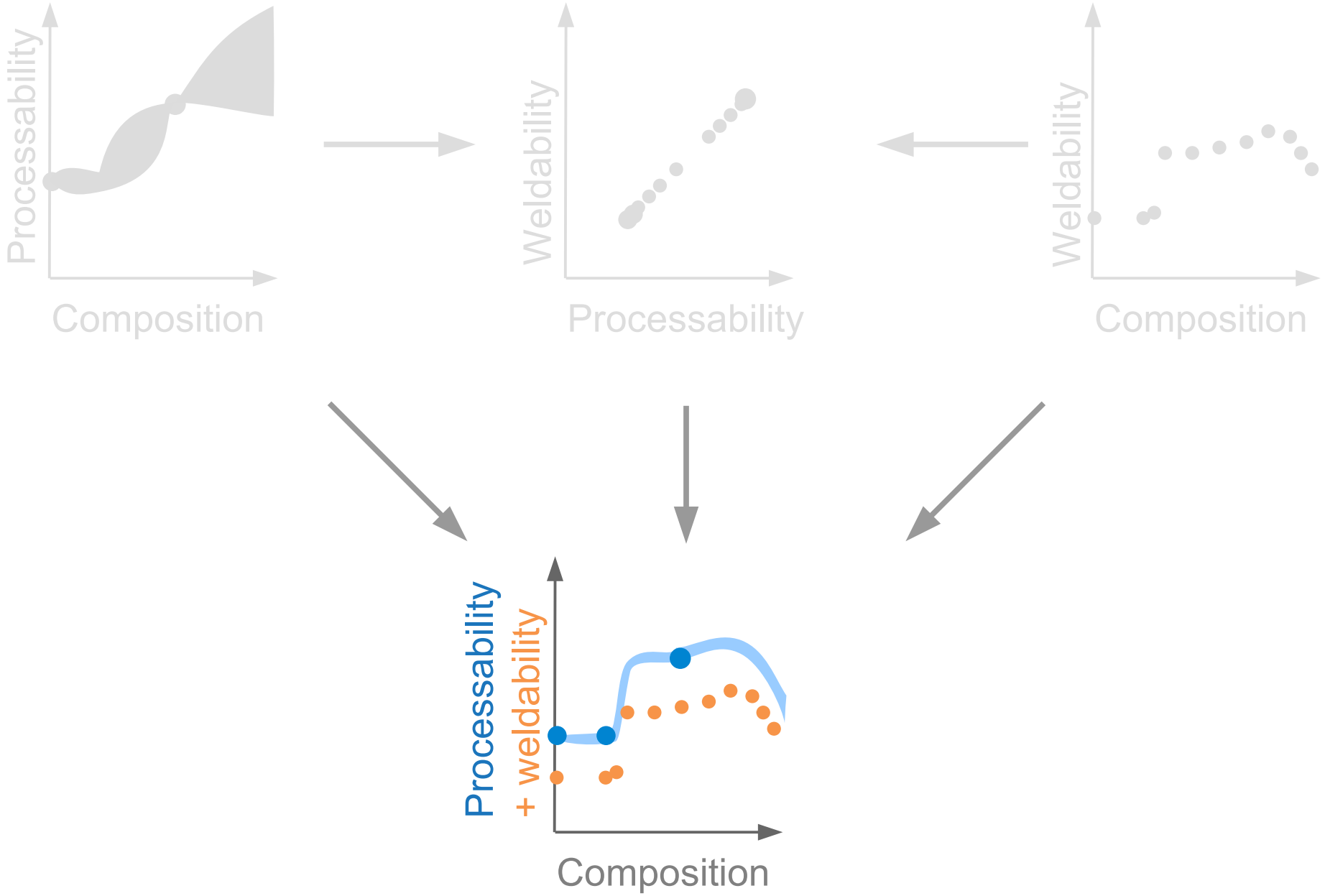


# Simple processability-welding relationship

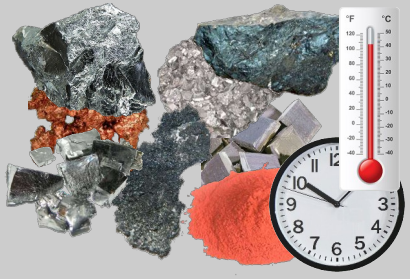


# Merging properties with the neural network



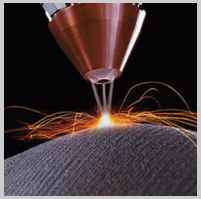
# Neural networks for materials design

## Composition



## Properties

Process



Fatigue

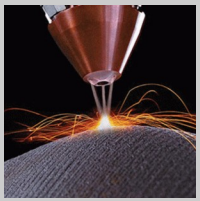
Welding

# Neural networks for materials design

**Composition**

**Properties**

Process



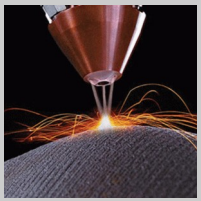
Fatigue

Welding

**Composition**

**Properties**

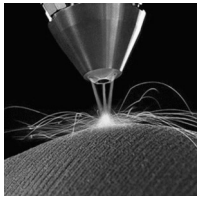
Process



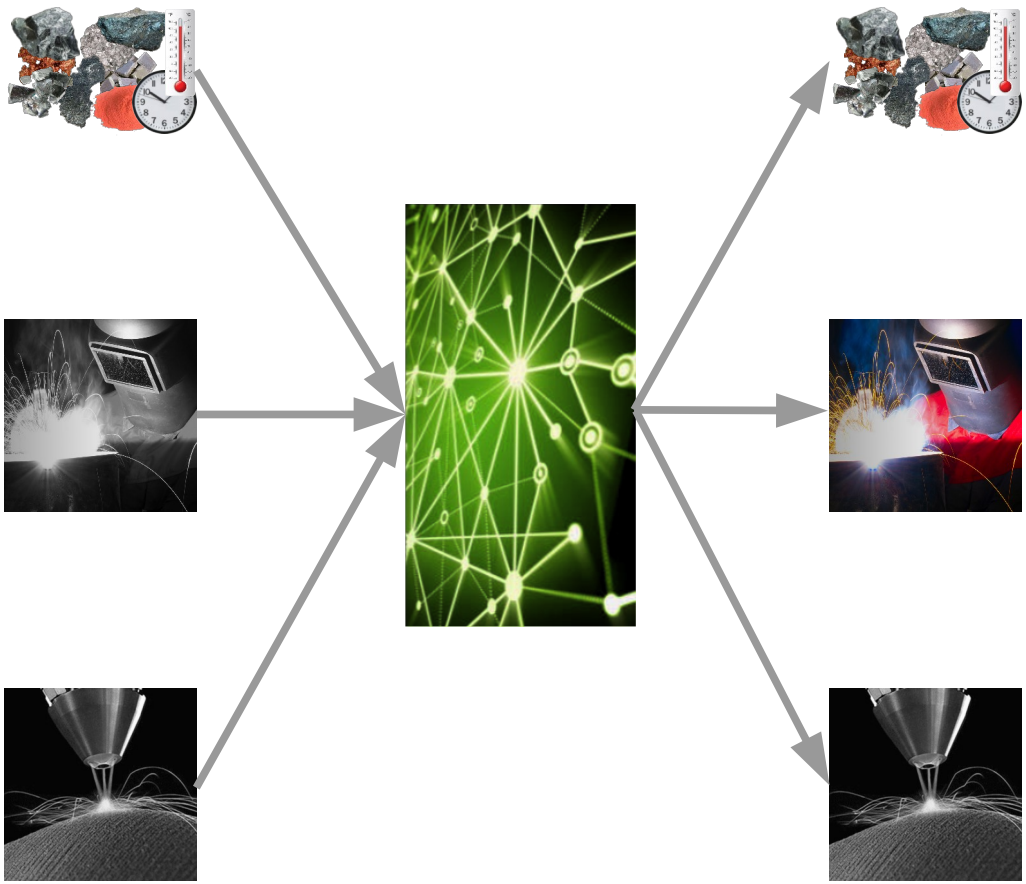
Fatigue

Welding

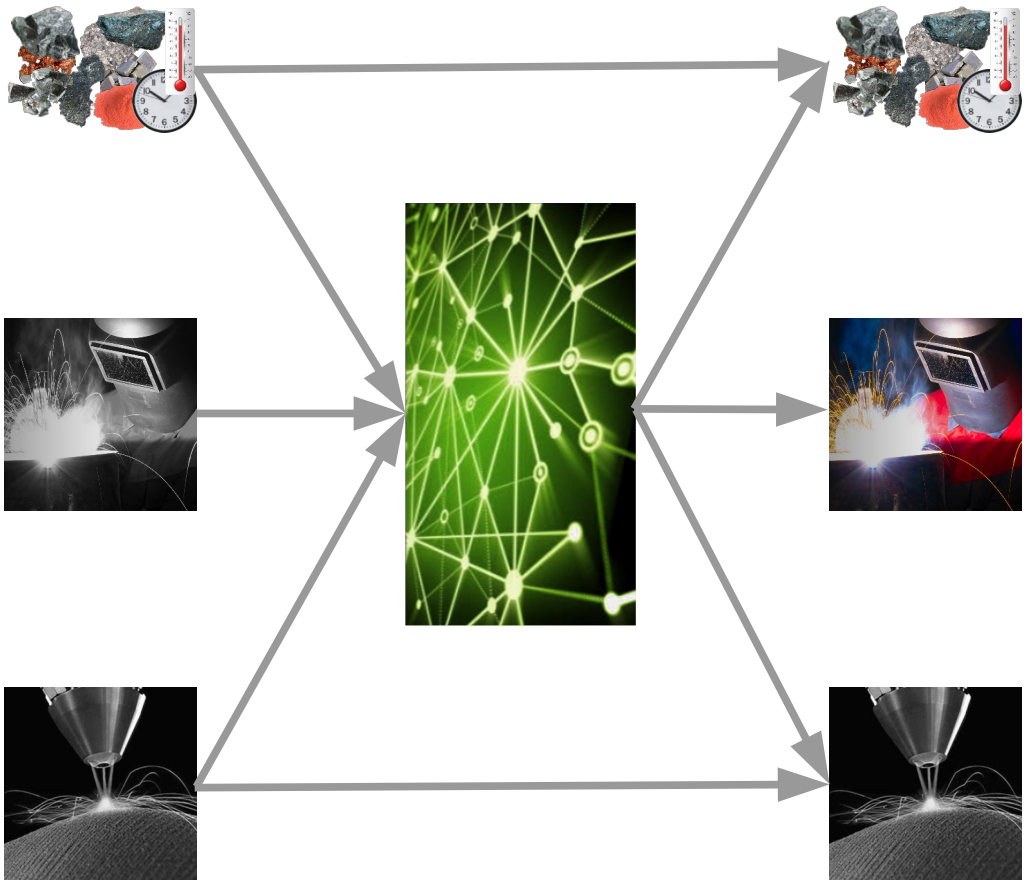
# Filling in missing values



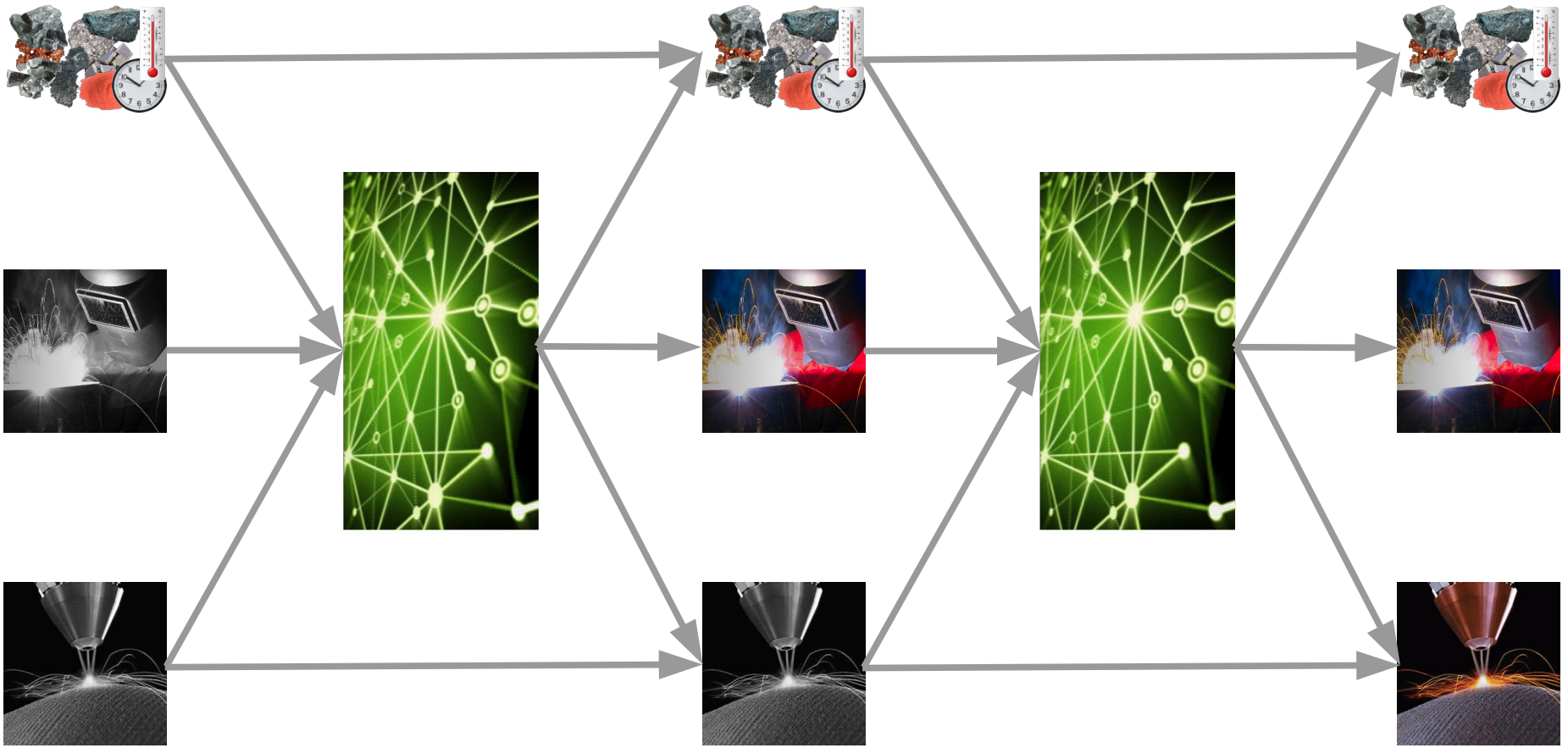
# Filling in missing values



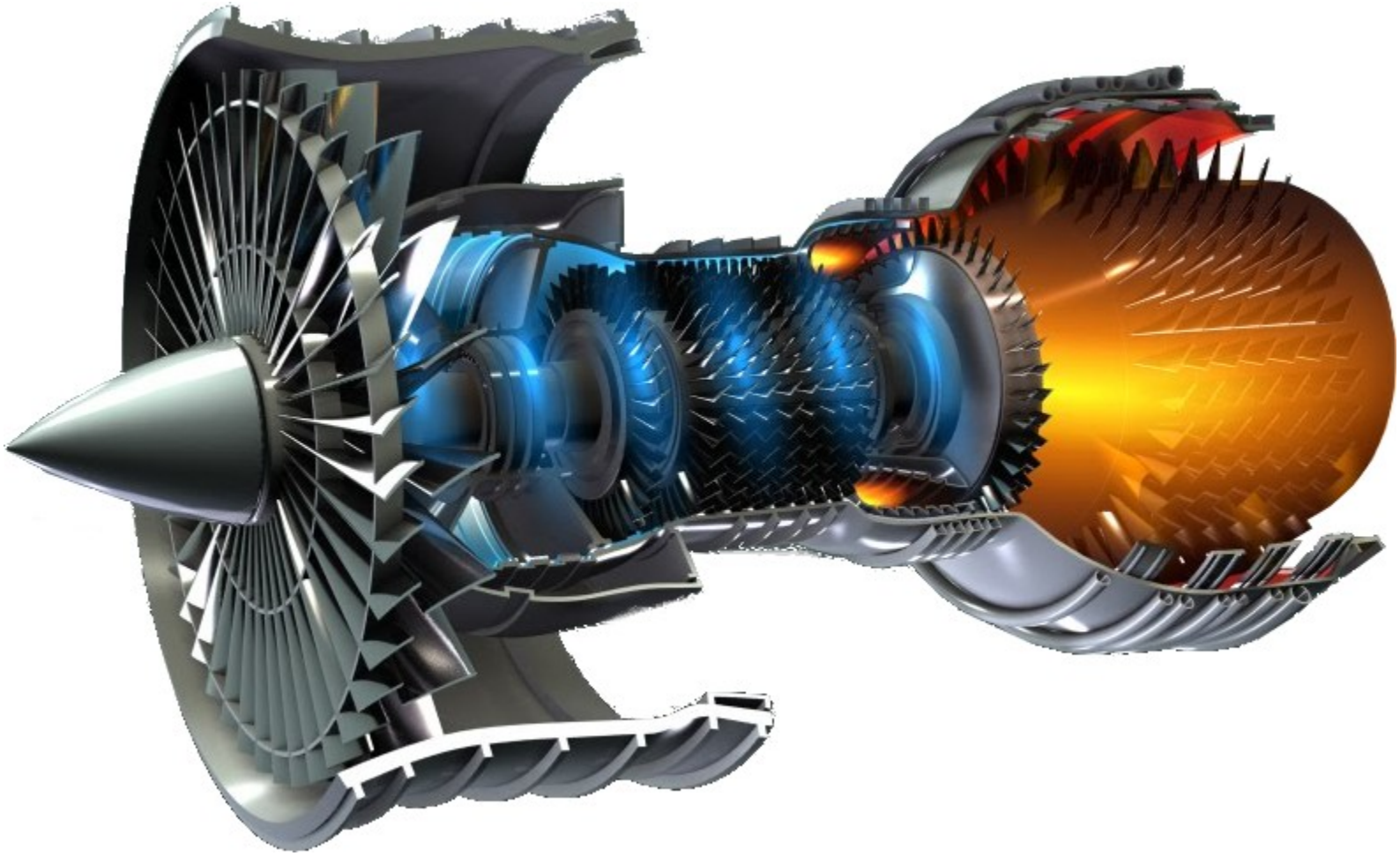
# Filling in missing values



# Filling in missing values



# Schematic of a jet engine



# Target properties

| | |
|---|---|
| Elemental cost | < 25 $\$kg^{-1}$ |
| Density | < 8500 $kgm^{-3}$ |
| γ' content | < 25 wt% |
| Oxidation resistance | < 0.3 $mgcm^{-2}$ |
| Processability | < 0.15% defects |
| Phase stability | > 99.0 wt% |
| γ' solvus | > 1000˚C |
| Thermal resistance | > 0.04 $K\Omega^{-1}m^{-3}$ |
| Yield stress at 900˚C | > 200 MPa |
| Tensile strength at 900˚C | > 300 MPa |
| Tensile elongation at 700˚C | > 8% |
| 1000hr stress rupture at 800˚C | > 100 MPa |
| Fatigue life at 500 MPa, 700˚C | > $10^5$ cycles |

# Composition

Cr: 19%

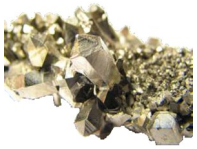
Co: 4%

Mo: 4.9%

W: 1.2%

Zr: 0.05%

Nb: 3%



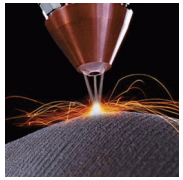
Al: 2.9%

C: 0.04%

B: 0.01%

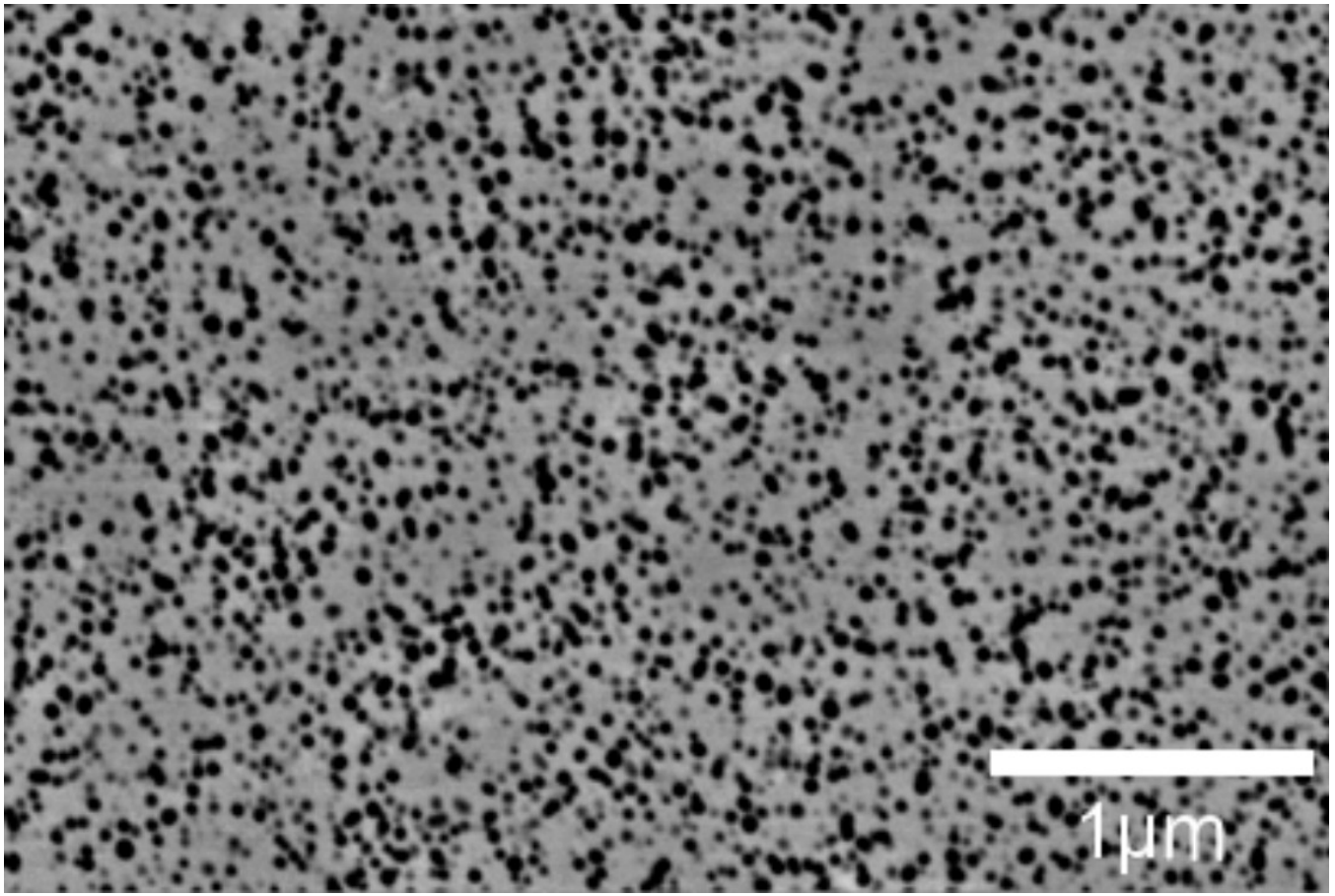
Ni

Expose 0.8

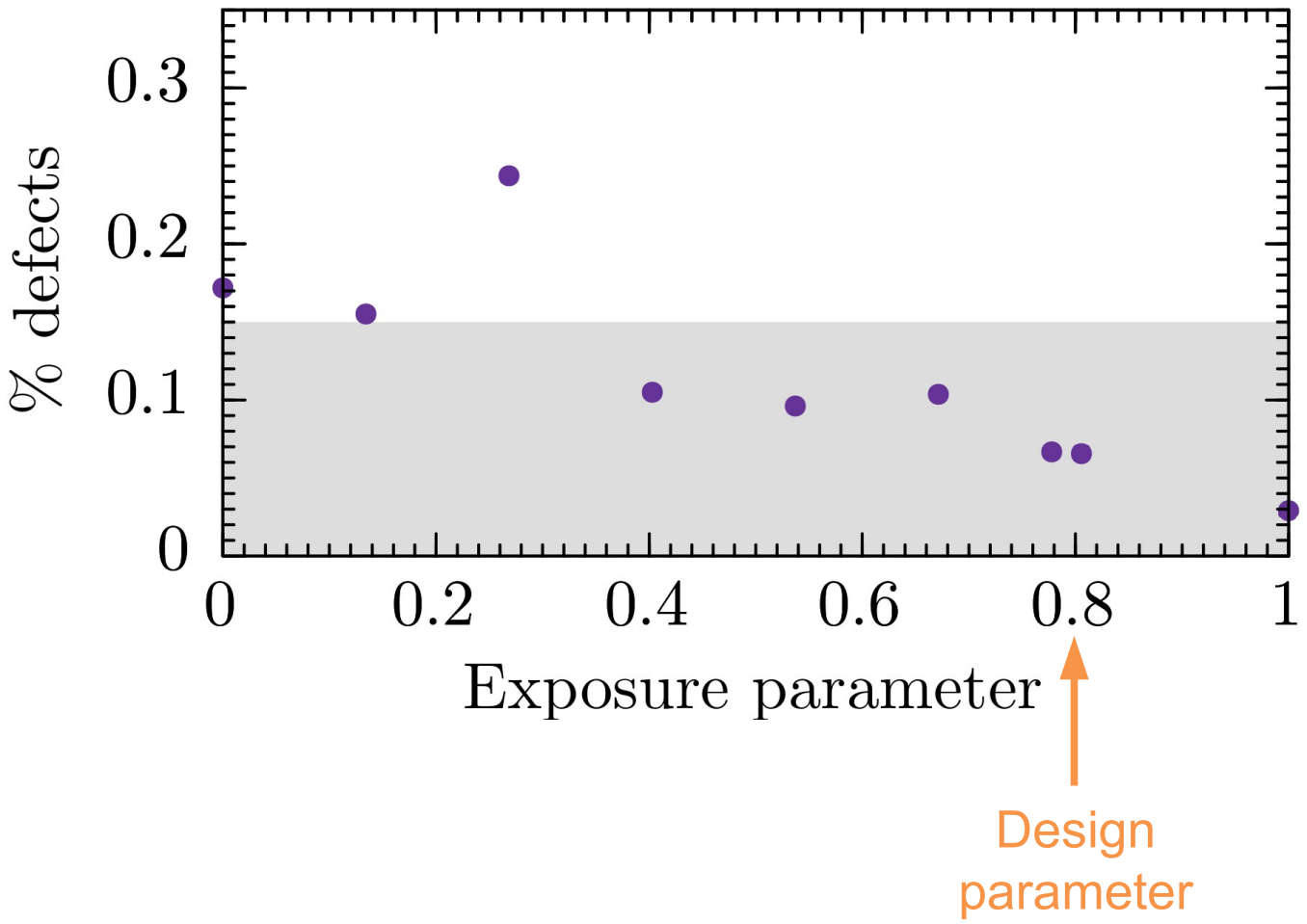


$T_{HT}$ 1300ºC

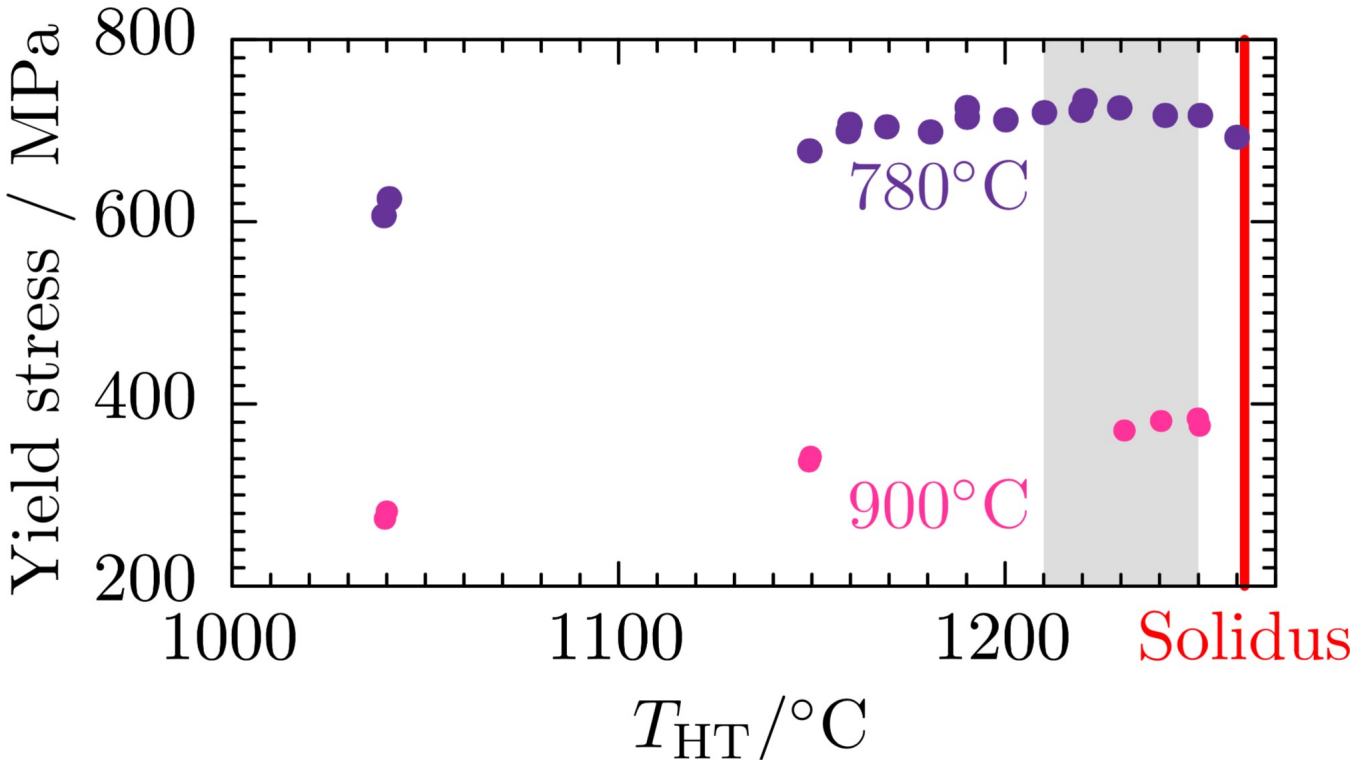
# Microstructure



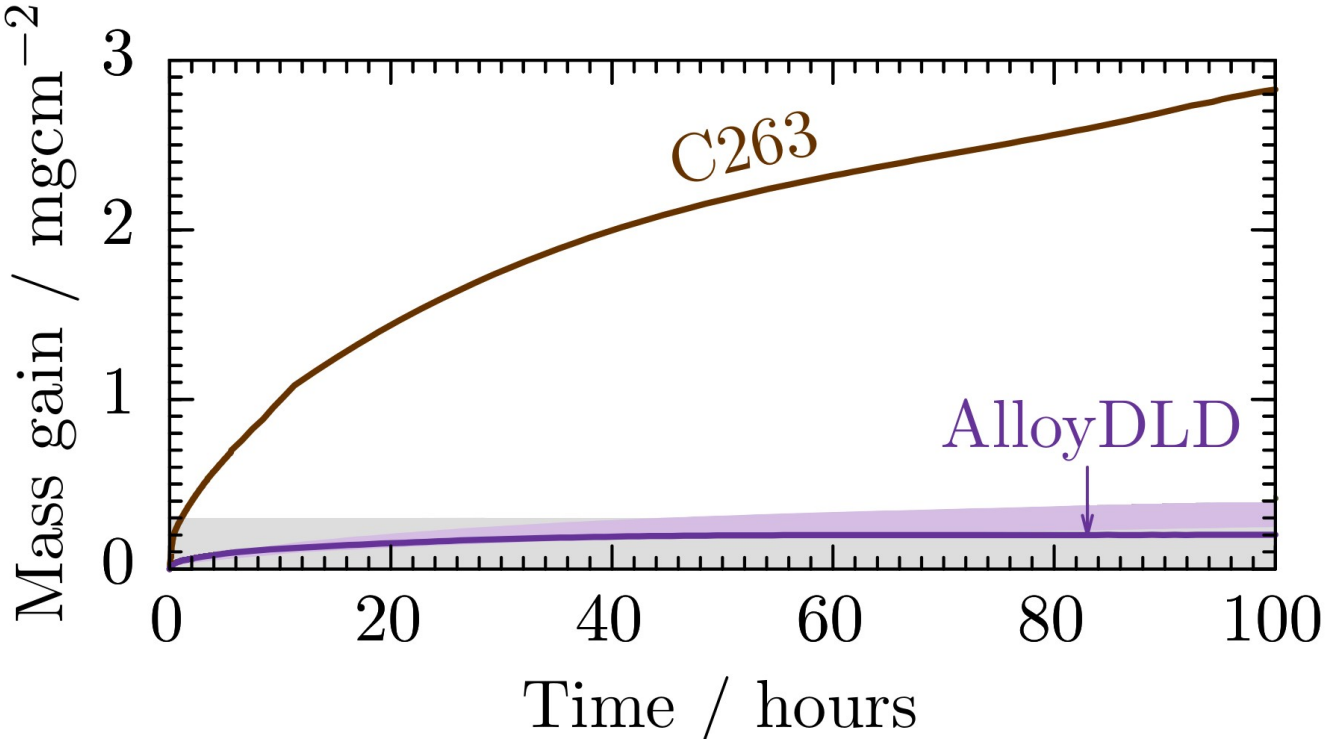
# Testing the processability: horizontal printing



# Testing the processability: horizontal printing



# Testing the oxidation resistance



# Printing components for an engine

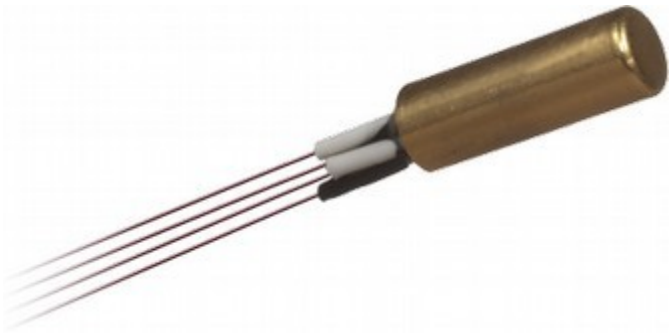
Materials Solutions

ROLLS ROYCE

# Low temperature physics

# Specification for a thermometer

90% of the cost of a thermometer is for calibration



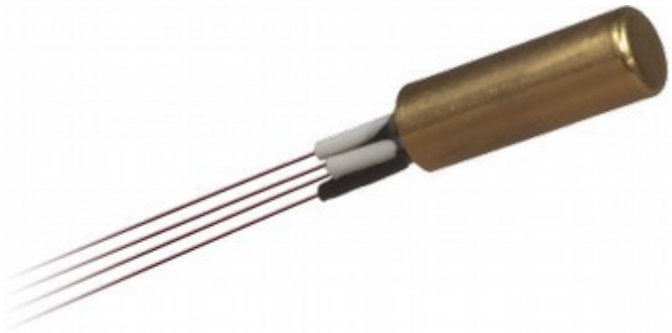
# Specification for a thermometer

90% of the cost of a thermometer is for calibration

Require a simple resistance-temperature relationship over a wide temperature range

Want constant sensitivity $T/R\ \mathrm{d}R/\mathrm{d}T$ with temperature

Thermometer must be stable with time and temperature



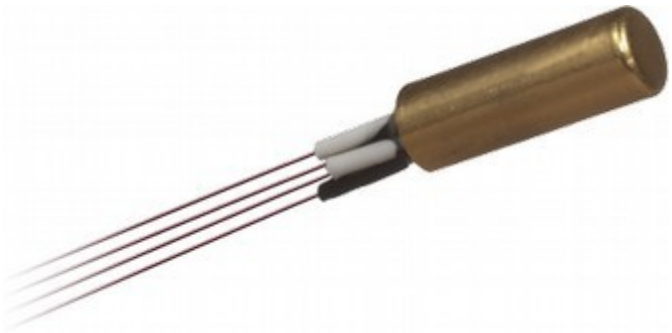
# Specification for a thermometer

15 experiments

100 DFT simulations

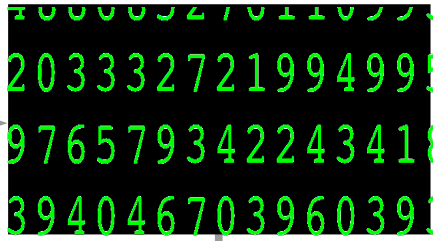
20000 CALPHAD calculations

100000 analytical results from quantum theory

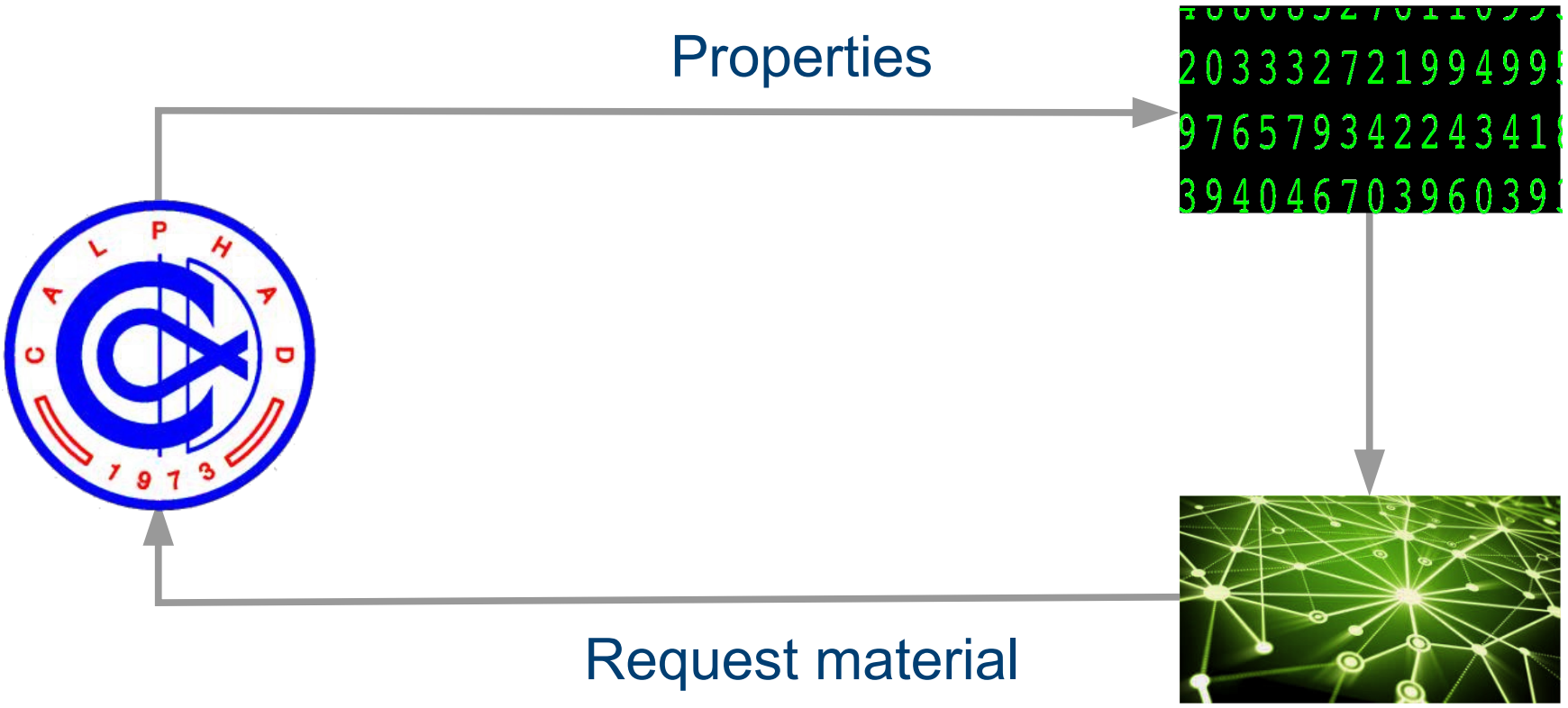


# Flowchart to train neural network

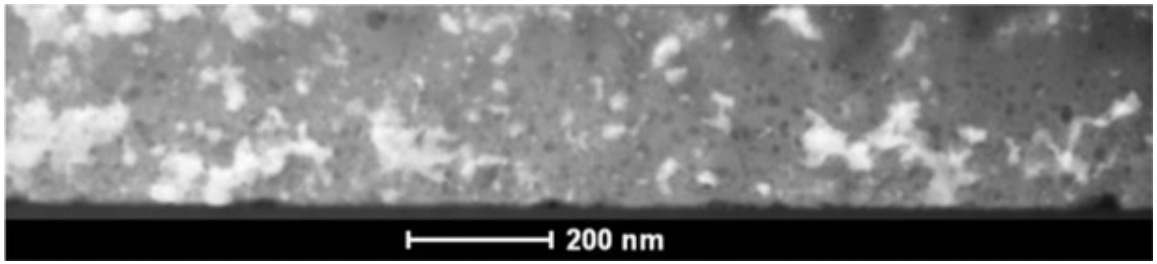

Properties



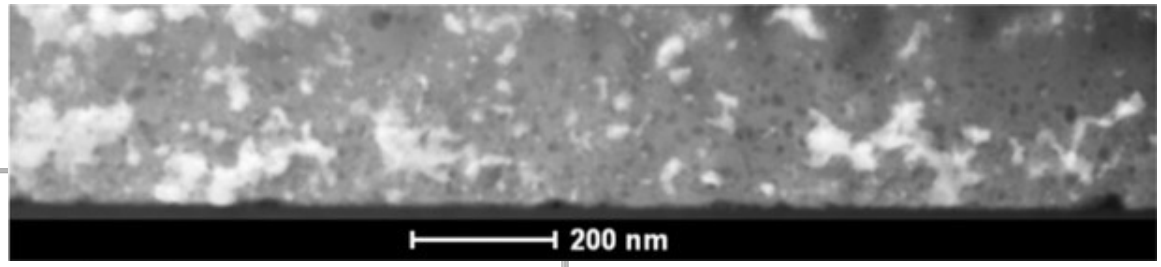
# Flowchart with reinforcement learning

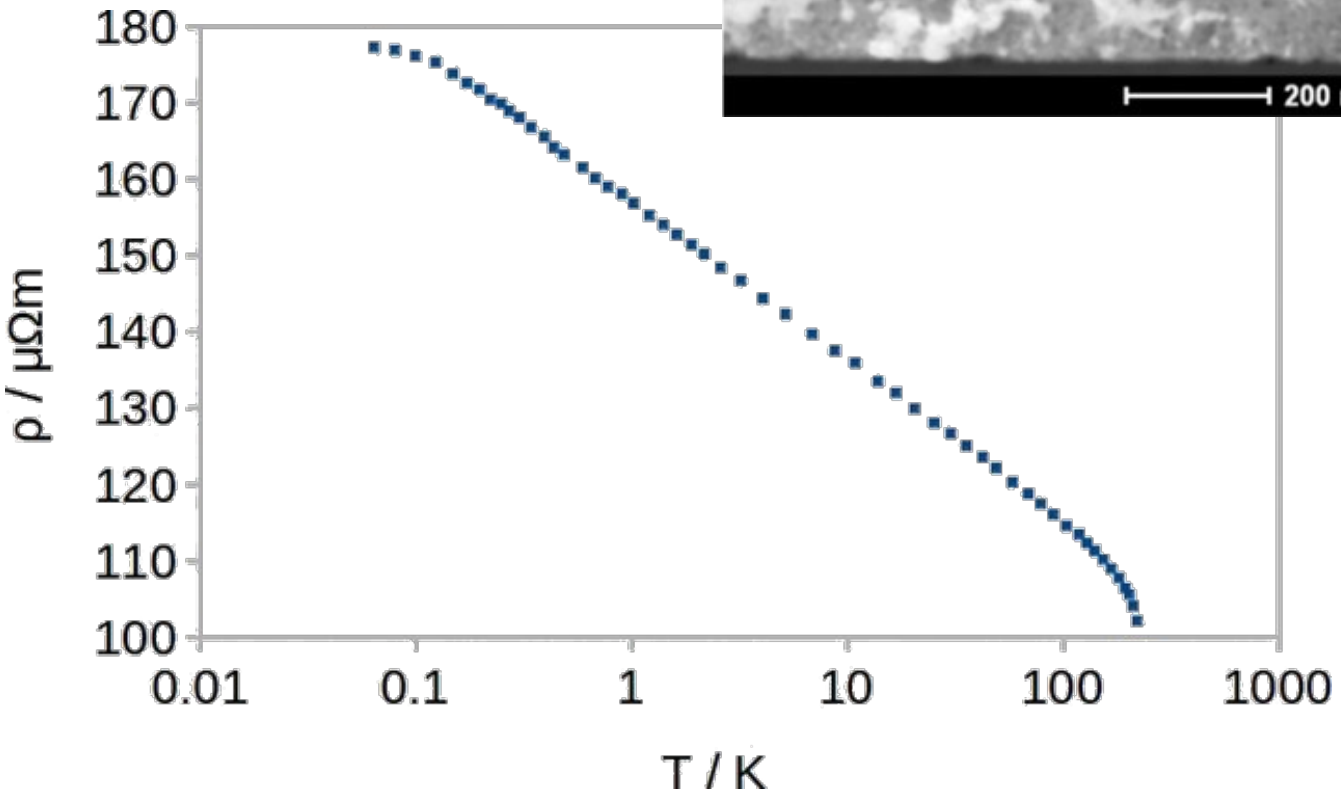


# Experimental verification of the thermometer

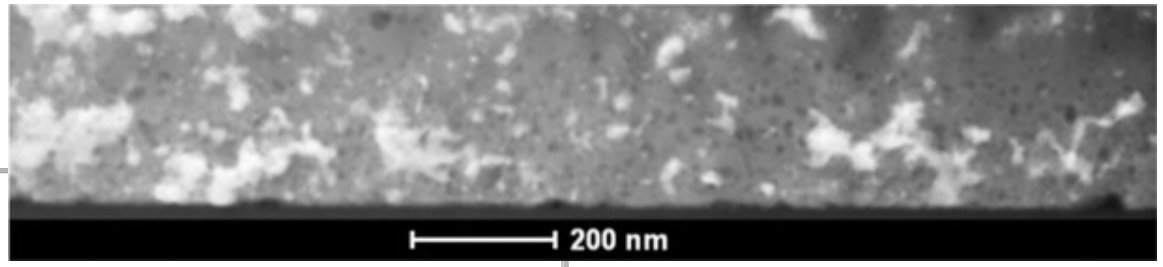


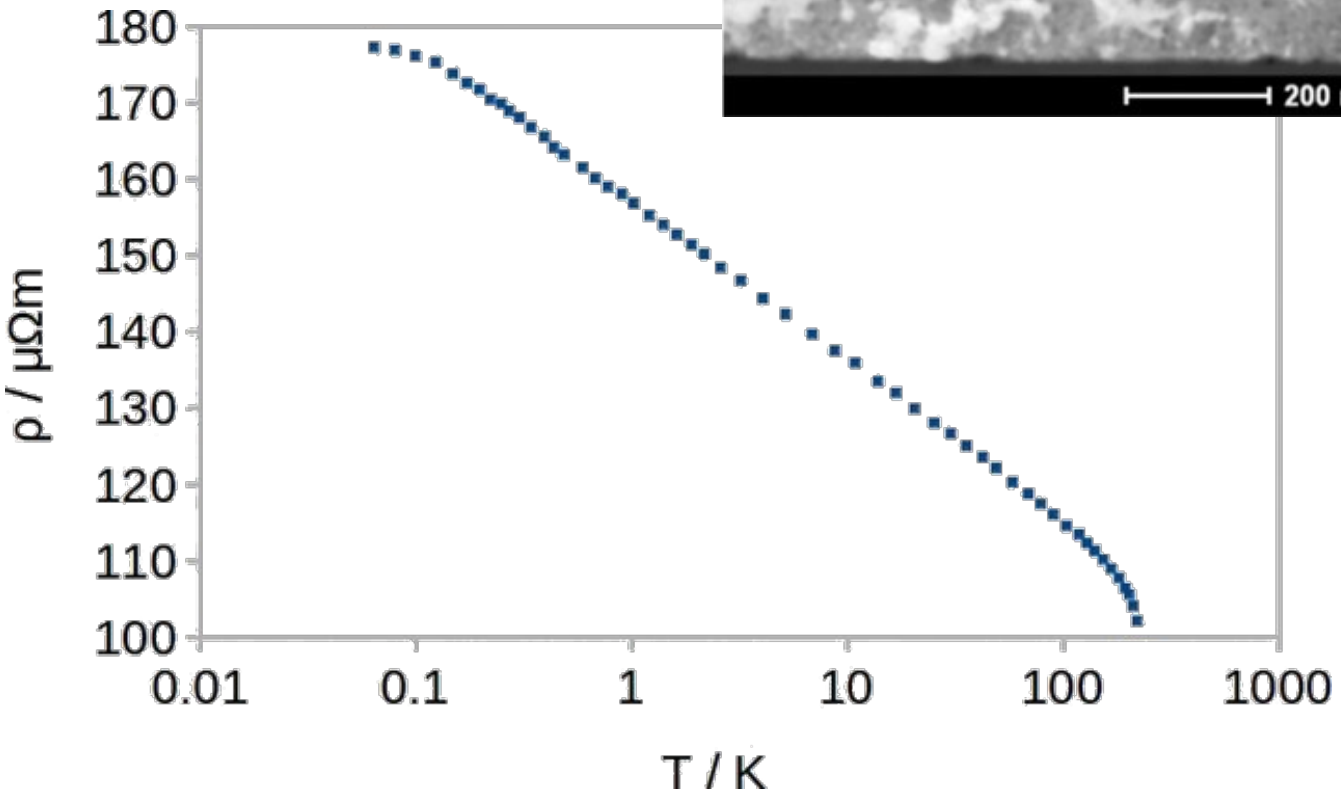
# Experimental verification of the thermometer



# Experimental verification of the thermometer



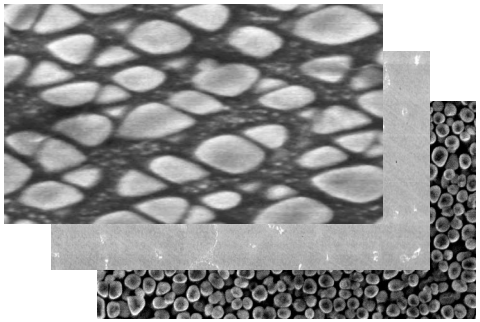




Measurements stable over 25 cycles and 6 months

Thermometer being sold by Cambridge Cryogenics

# Materials designed

Nickel and molybdenum

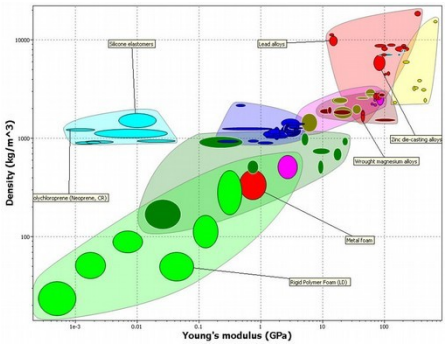


Experiment and DFT for batteries


Steel for welding

# More materials

Identified and corrected errors in materials database




GRANTA
MATERIAL INSPIRATION

Lubricants with molecular dynamics and experiments

bp

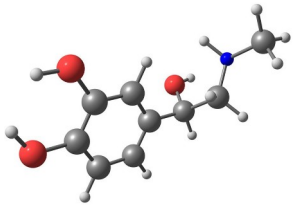
Drug design

e-therapeutics

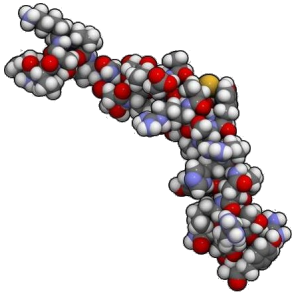
optibrium

Takeda

# Action of a drug



Drug

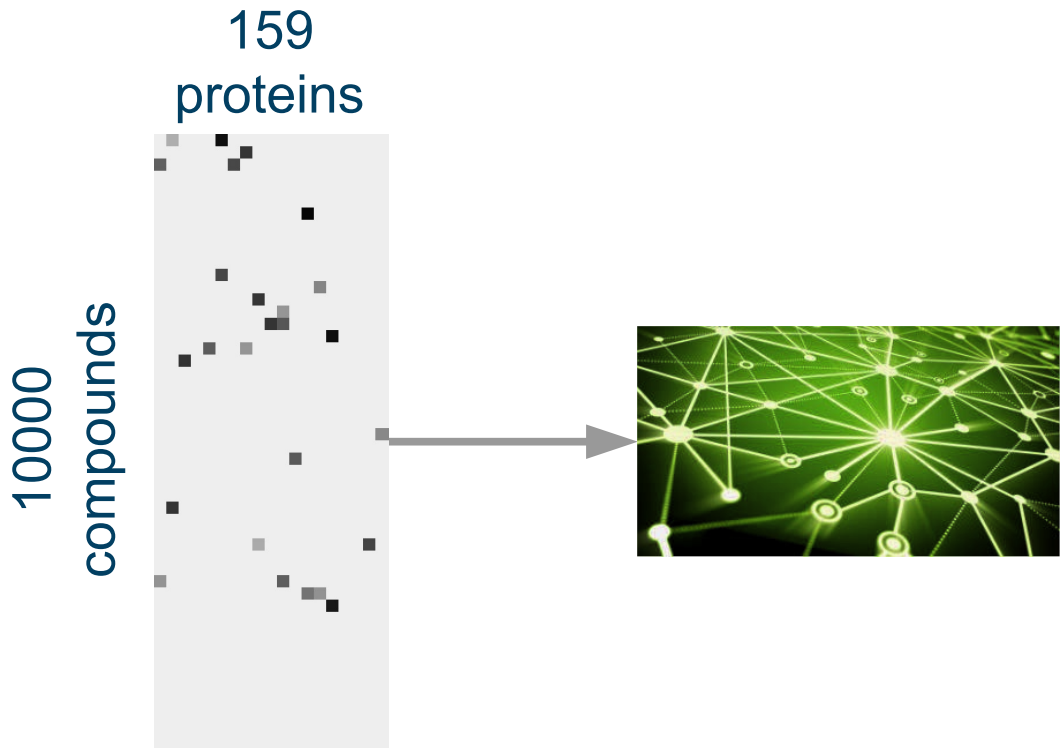


Protein

Effect

# Novartis dataset to benchmark machine learning

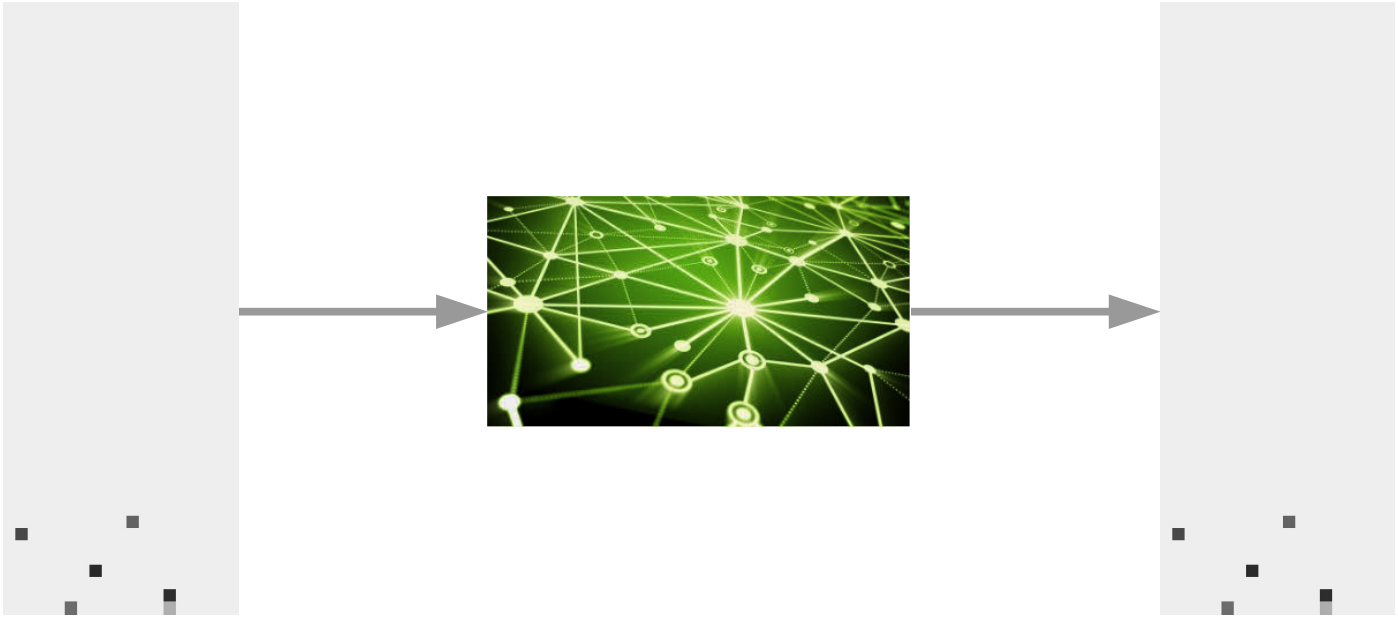
159 kinase proteins, 10000 compounds, data 5% complete



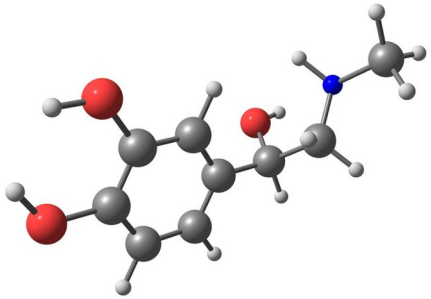

Data from ChEMBL
Martin, Polyakov, Tian, and Perez,
J. Chem. Inf. Model. 57, 2077 (2017)

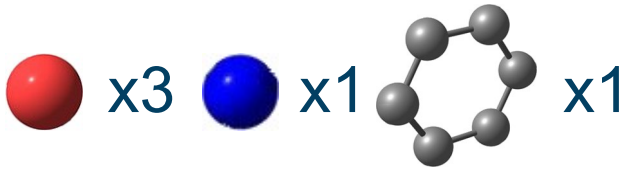
# Impute missing entries to validate

Validate using a realistically split holdout data set, extrapolate to new chemical space

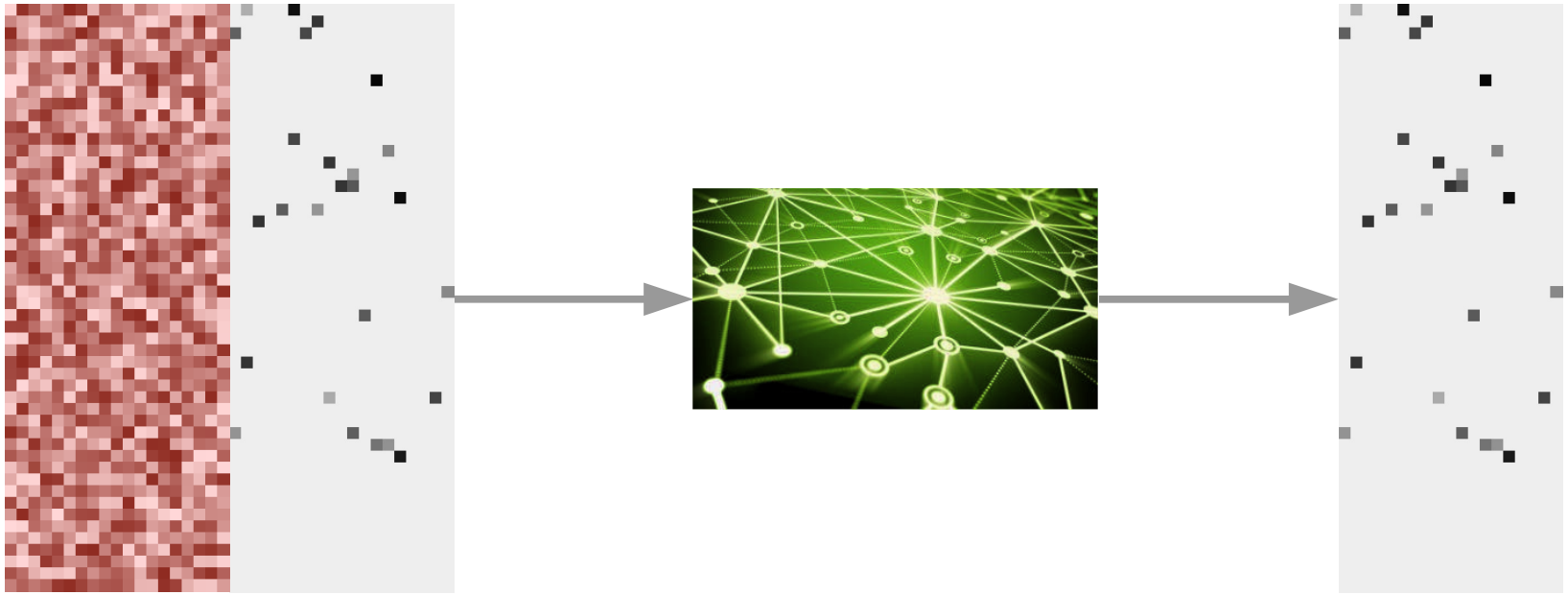


# Quantitative structure-activity relationships

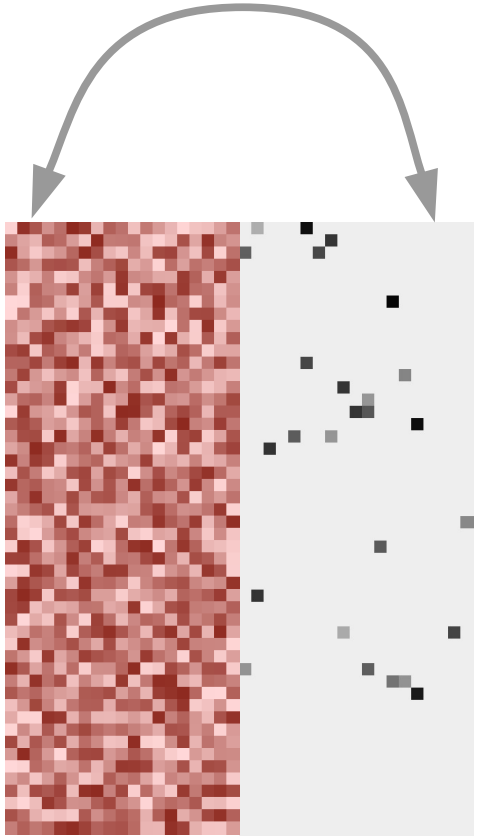





Molecular weight=183 Da

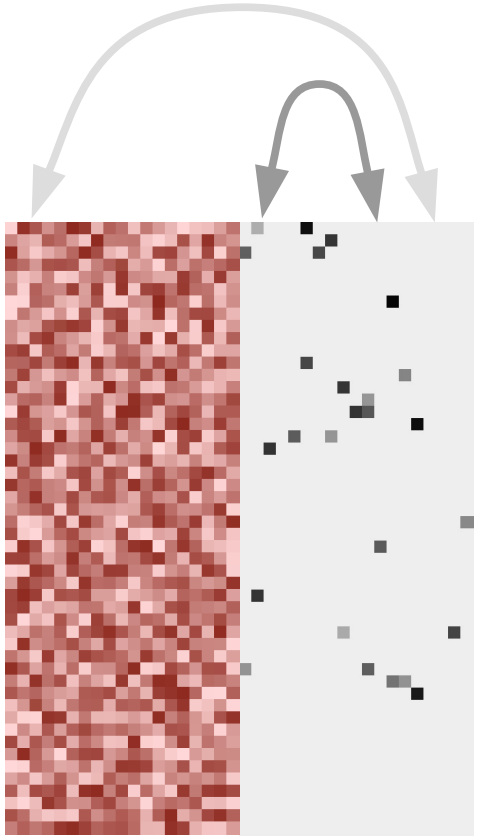


# Quantitative structure-activity relationships



Standard methods learn chemical descriptor-protein correlations

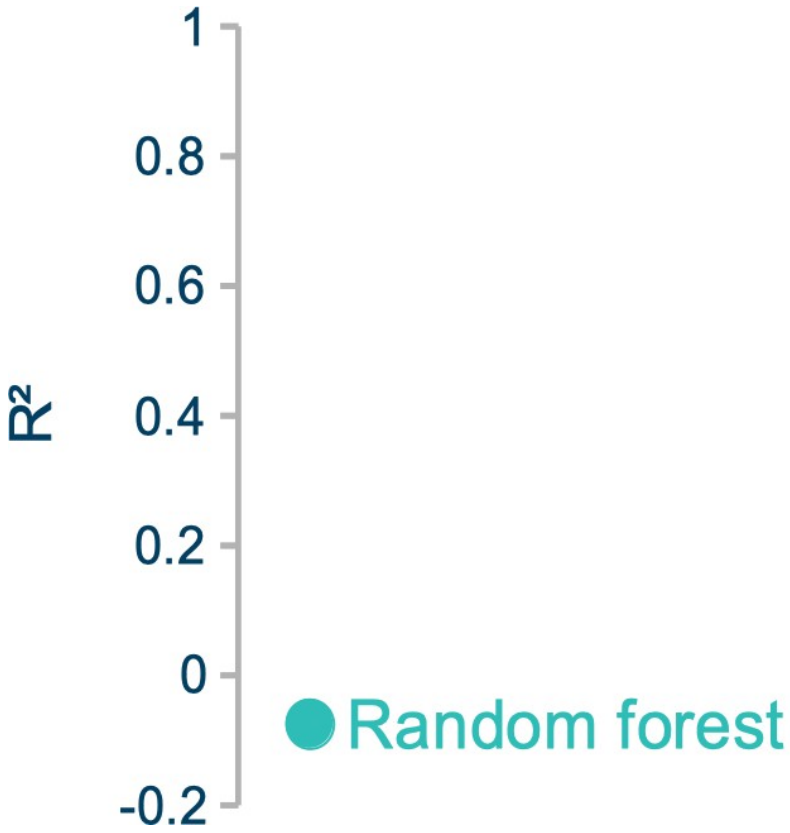
# Quantitative structure-activity relationships



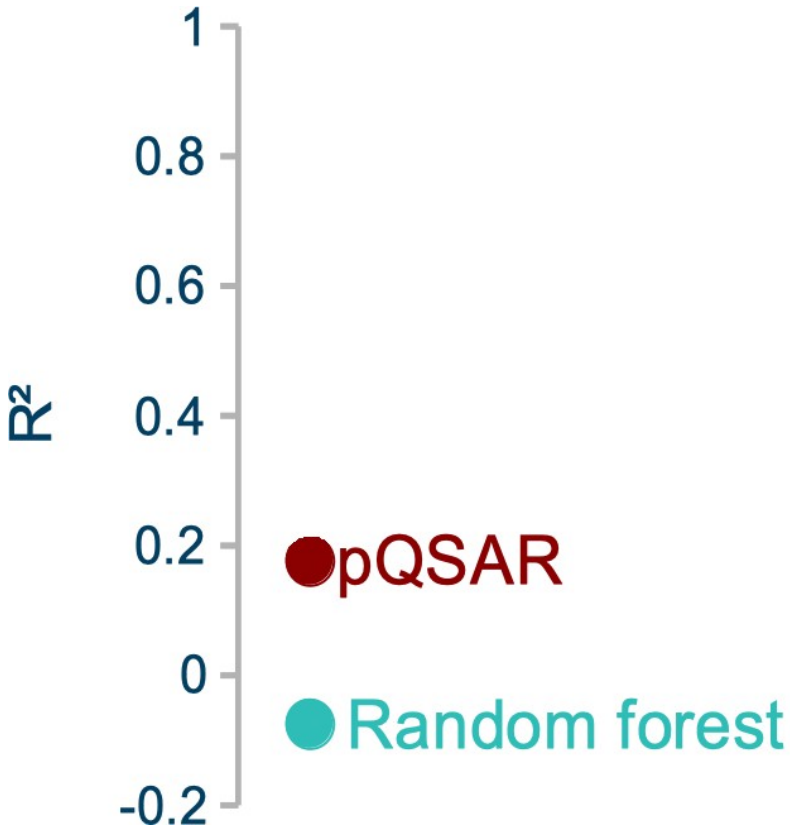
Standard methods learn chemical descriptor-protein correlations

Neural network also learns the protein-protein correlations

# Random forest

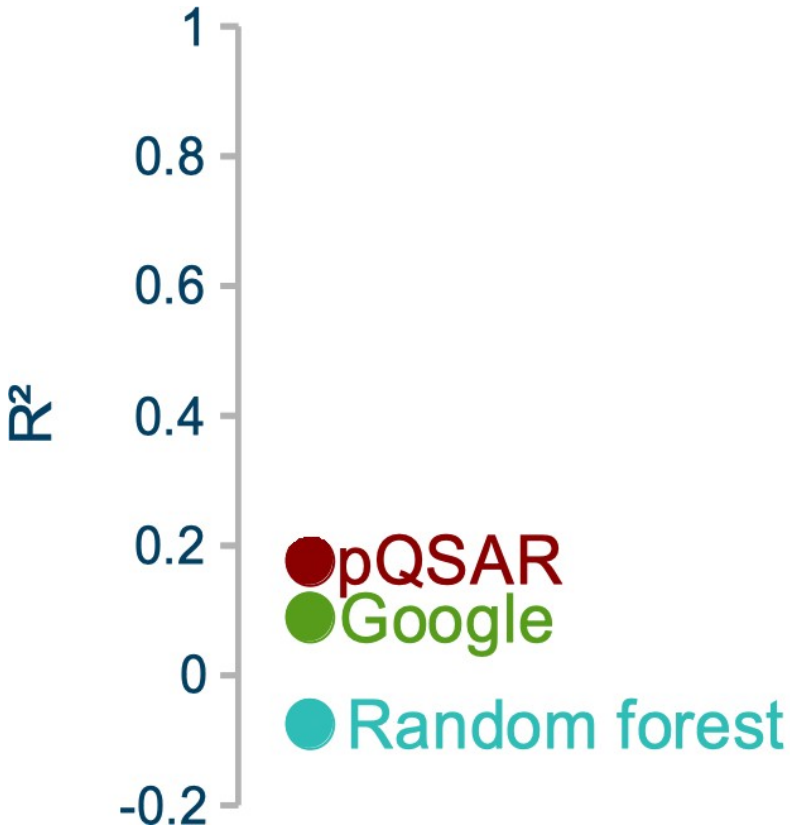


# Predictions from pQSAR

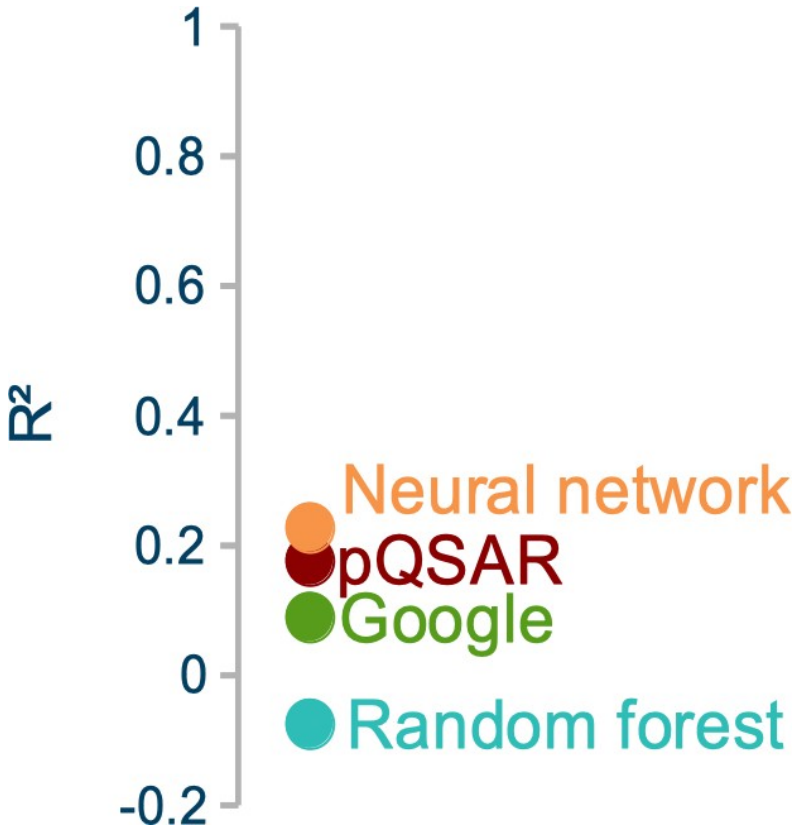



Martin, Polyakov, Tian, and Perez,
J. Chem. Inf. Model. 57, 2077 (2017)

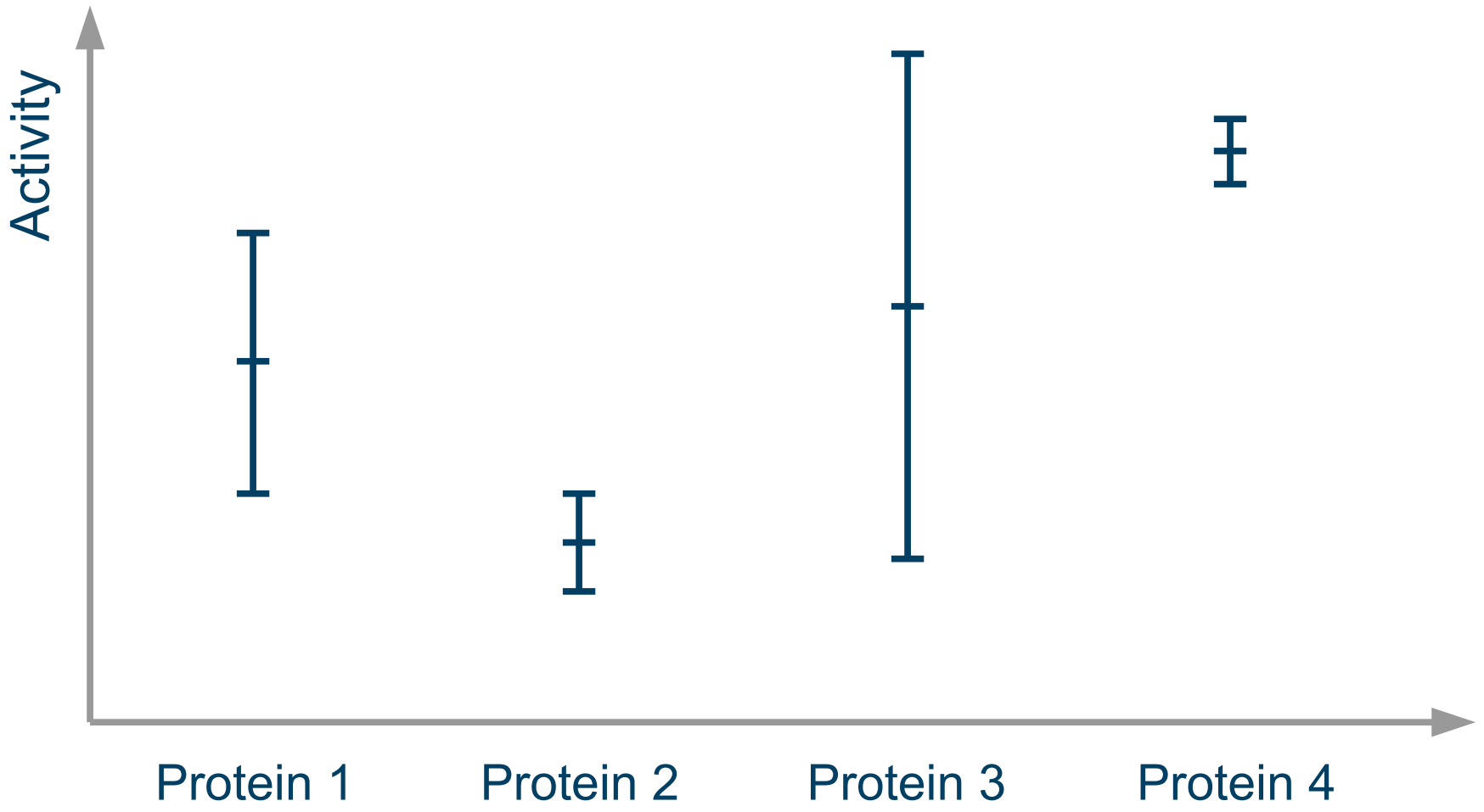
# Google’s attempt



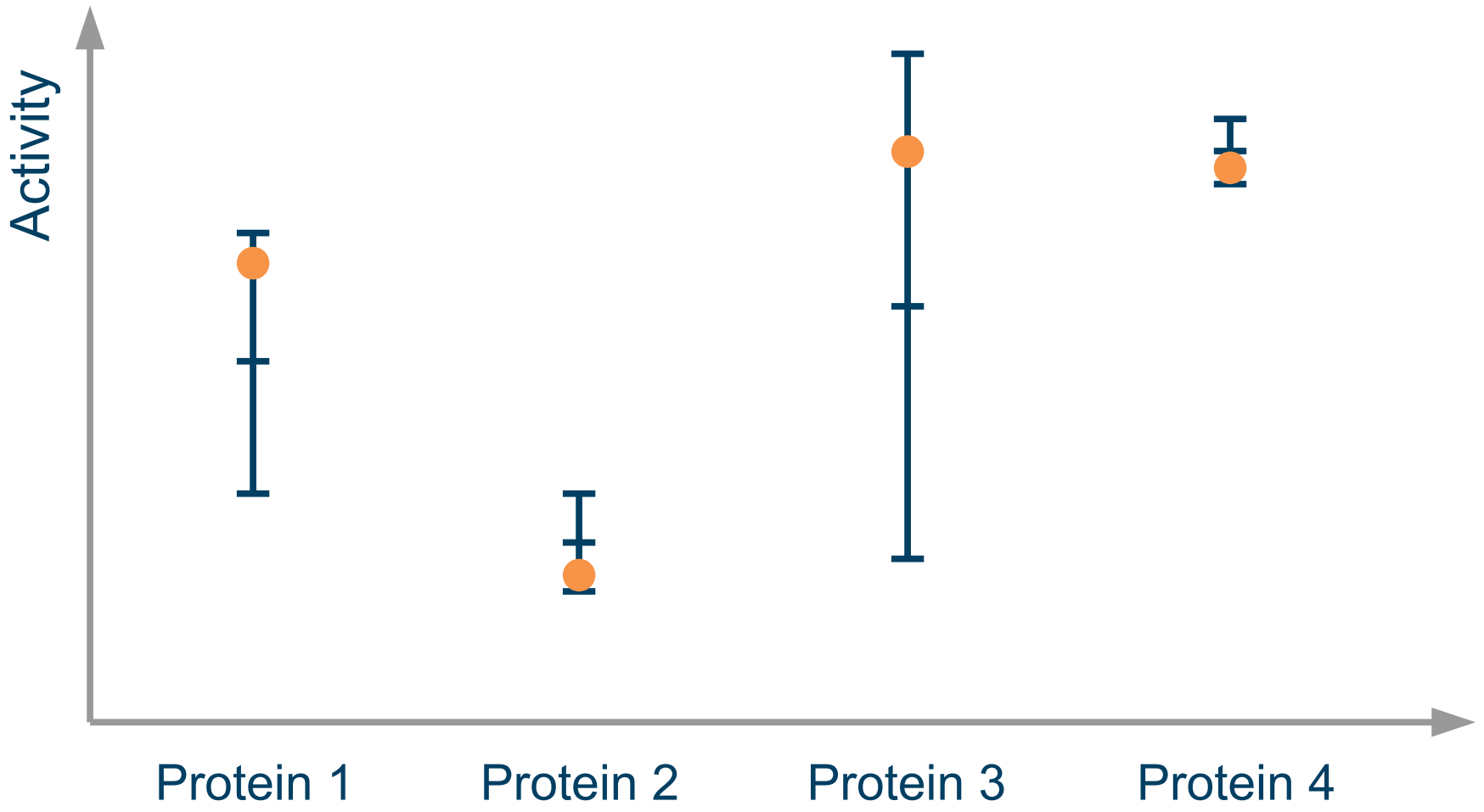
# Neural network with missing data



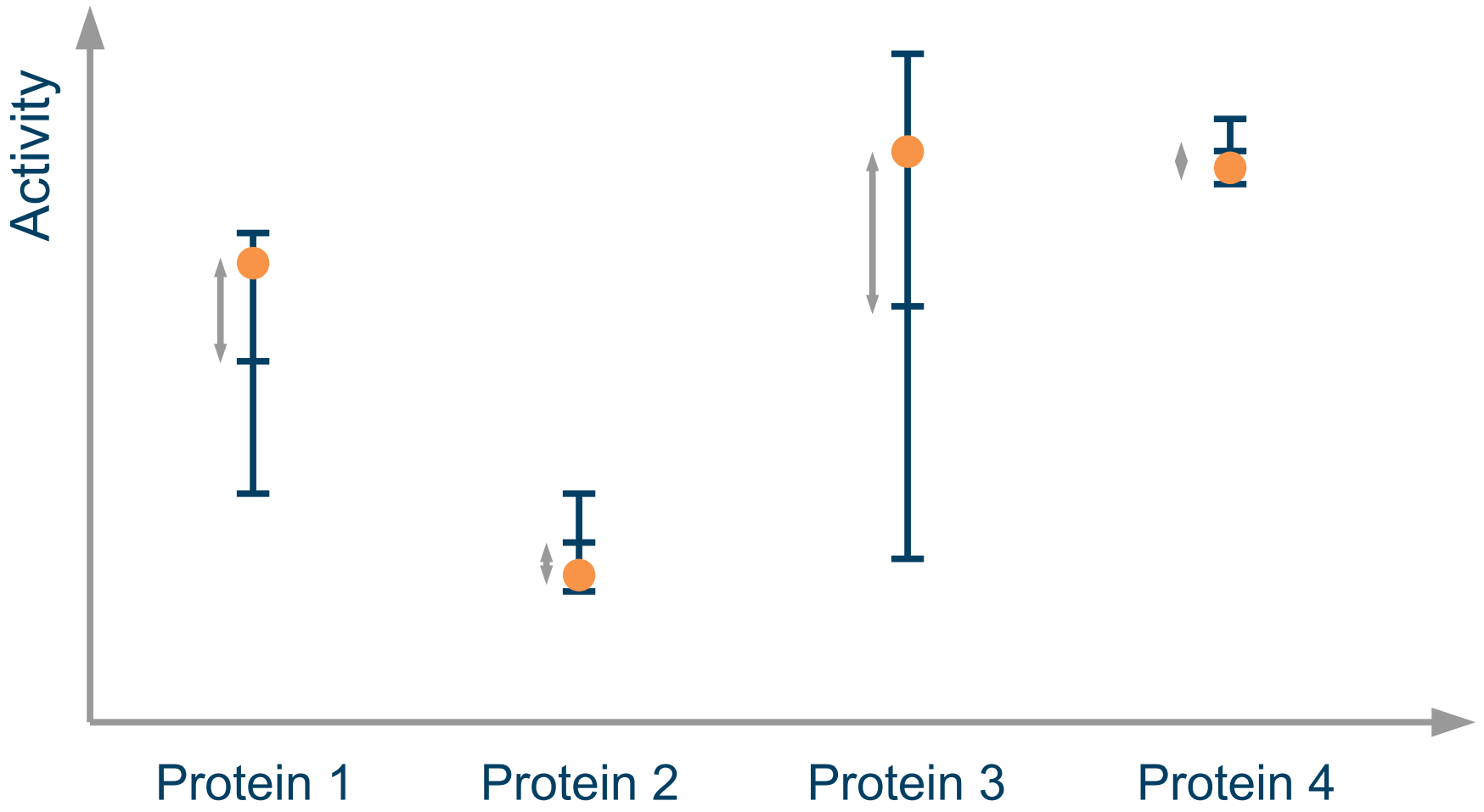
# Predictions have an uncertainty



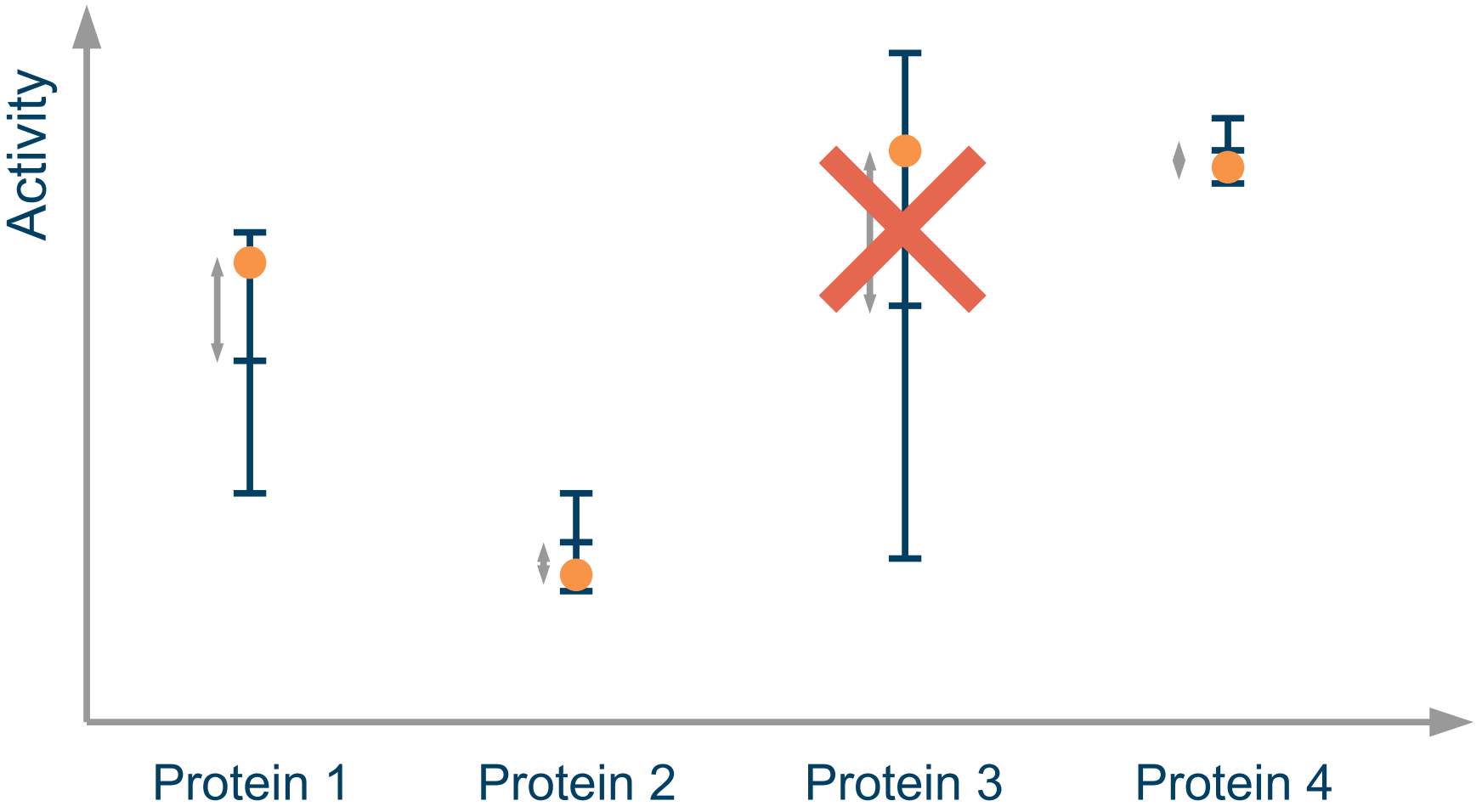
# Validation data typically within one standard deviation



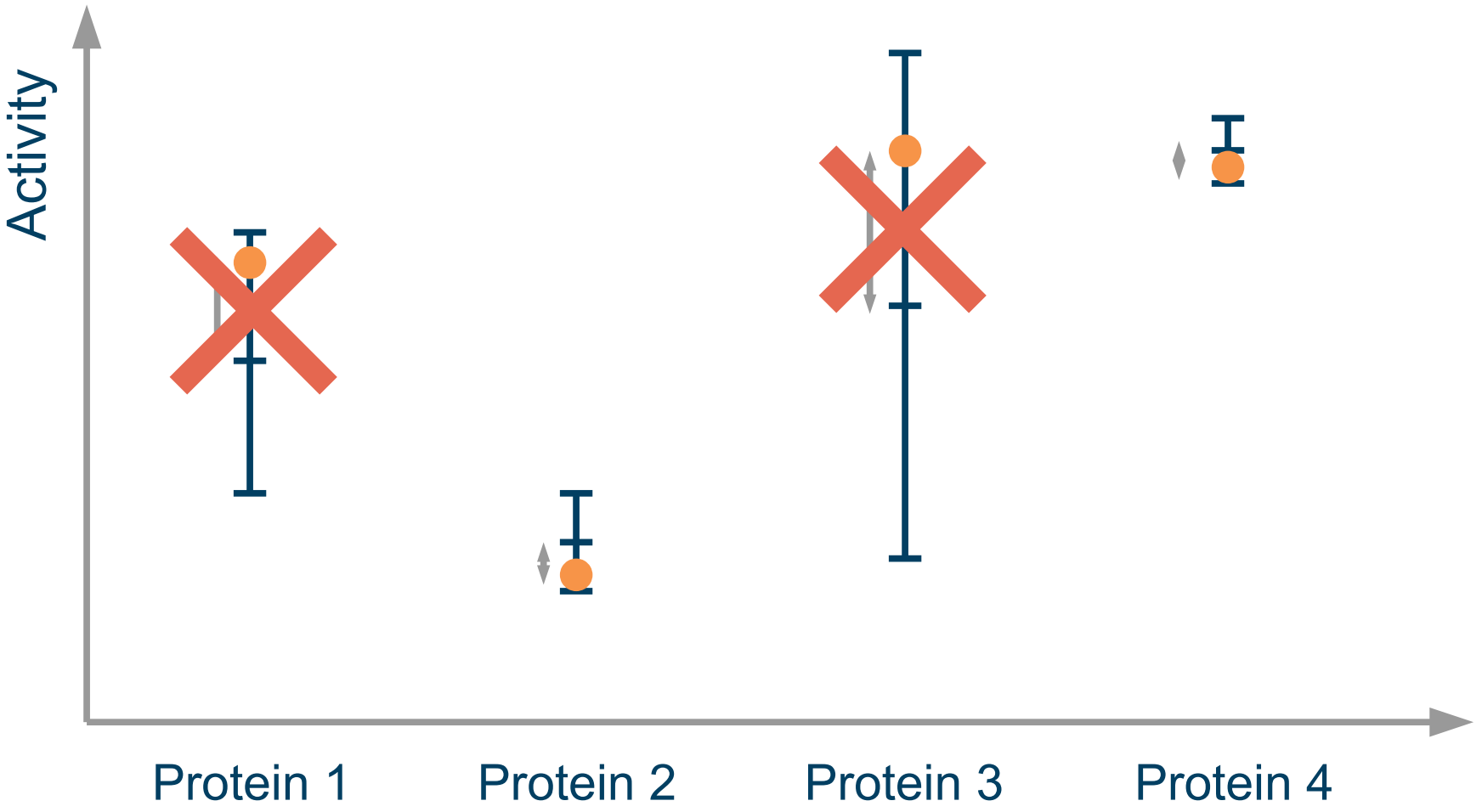
# $R^2$ metric calculated with difference from mean



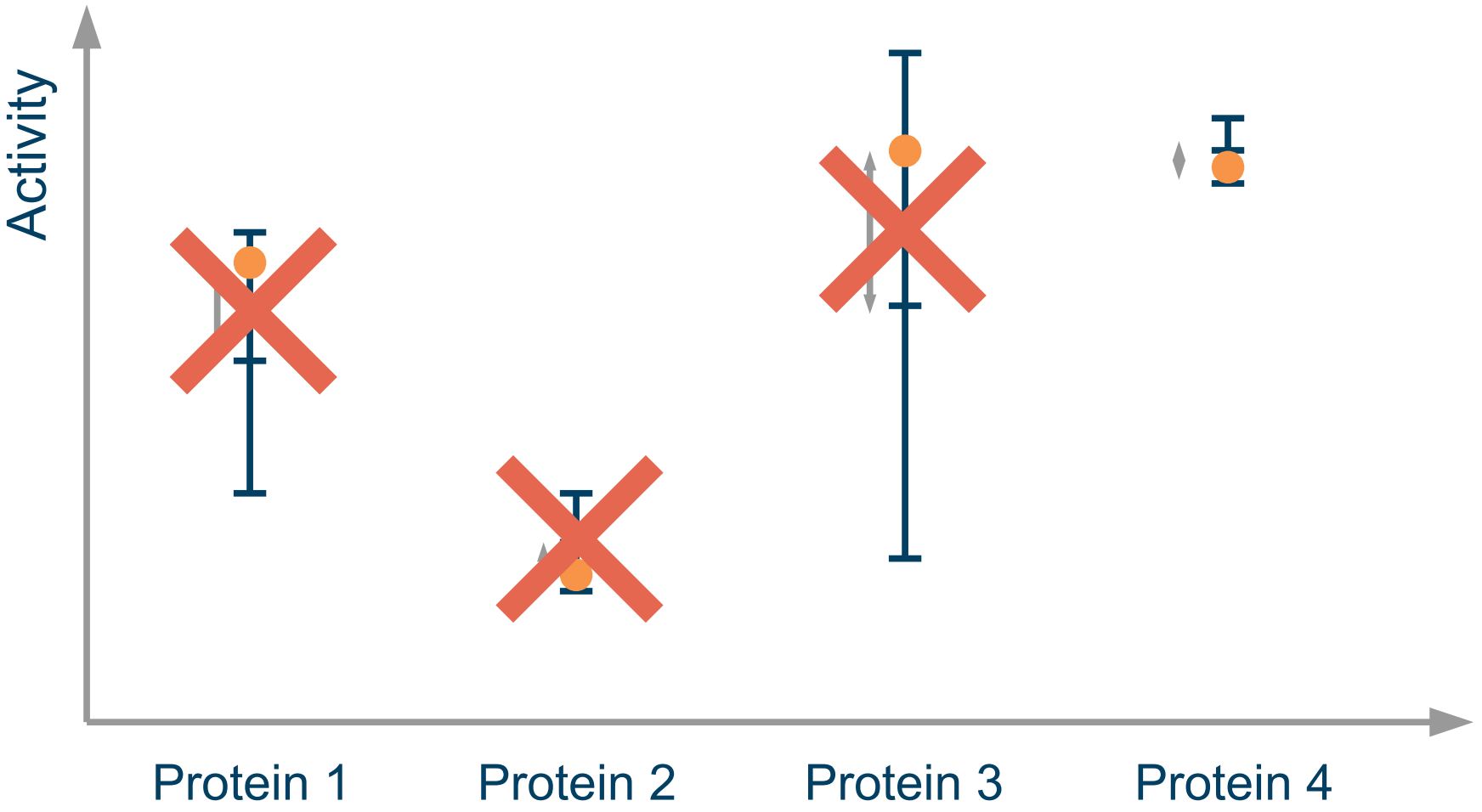
# Impute 75% of data with smallest uncertainty



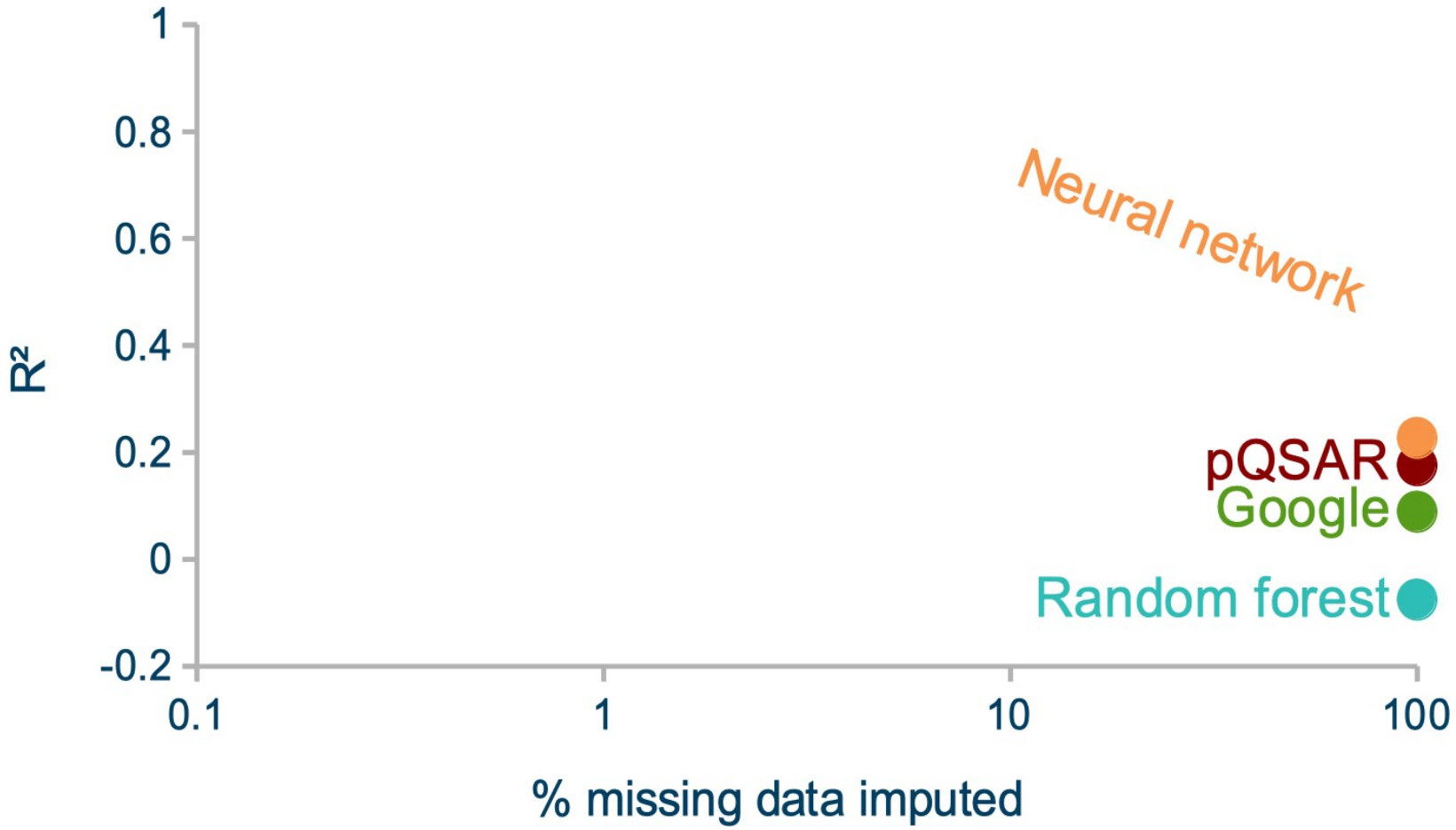
# Impute 50% of data with smallest uncertainty



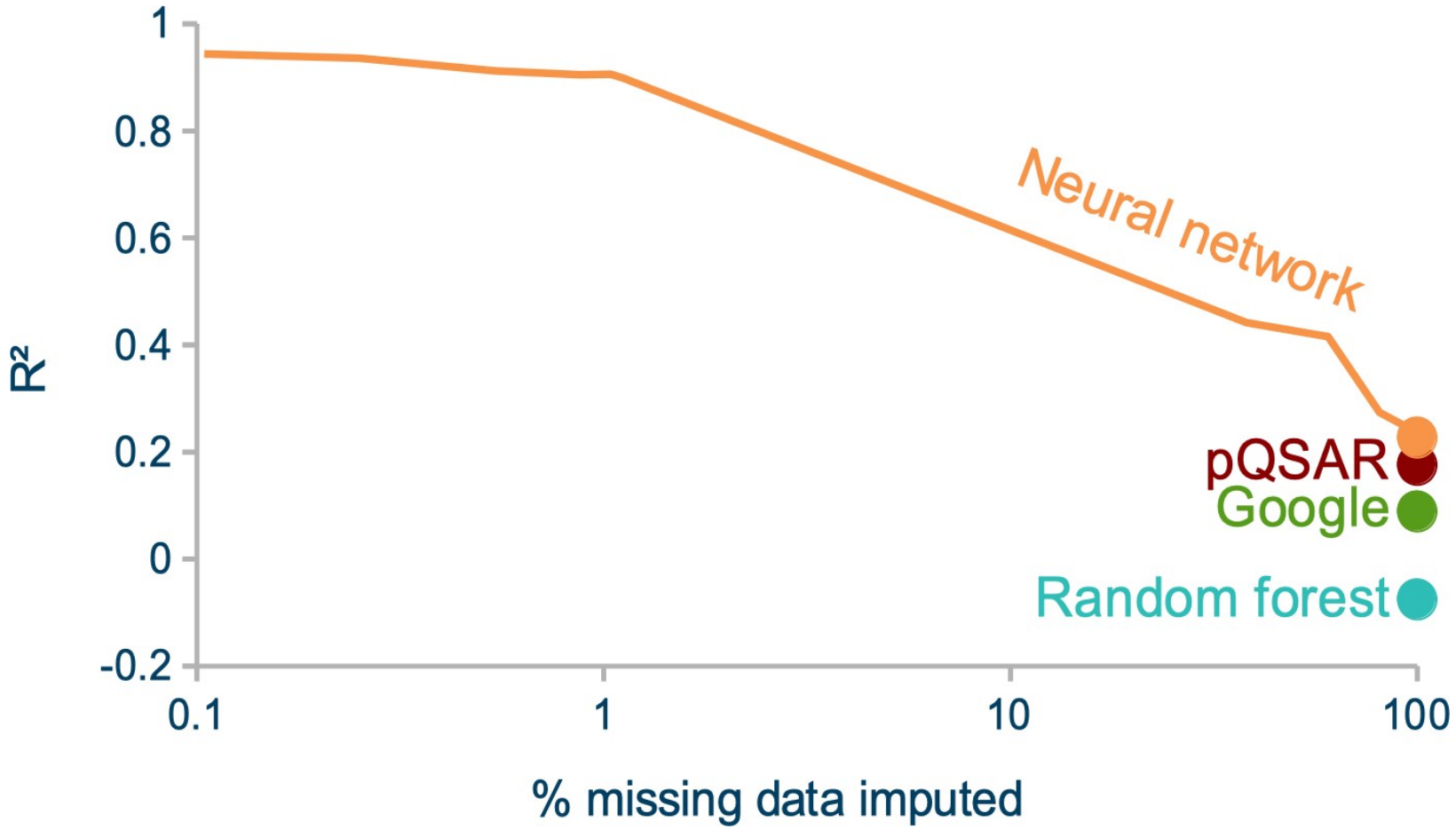
# Impute 25% of data with smallest uncertainty



# Improved performance by exploiting uncertainty

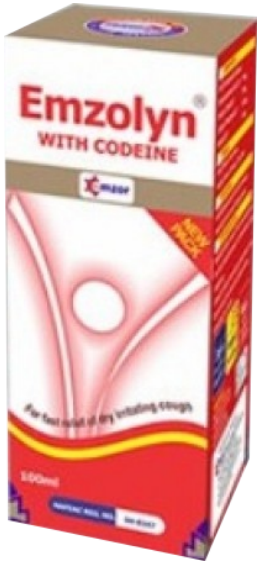


# Improved performance by exploiting uncertainty



# Different drugs can treat the same ailment



# Roadmap to productization

Reseller agreement with drug discovery software company optibrium

Machine learning tool embedded into next generation of optibrium software for release in October 2020

# Summary

Merge different experimental quantities and computer simulations into a holistic design tool

Designed and experimentally verified alloy for direct laser deposition

Improved predicability of drug design from $R^2$=0.18 to $R^2$=0.93

Additional experimentally proven materials, founded start-up intellegens.ai

https://app.intellegens.ai/steel_search