Materials discovery with artificial intelligence

Gareth Conduit

Neural network algorithm to

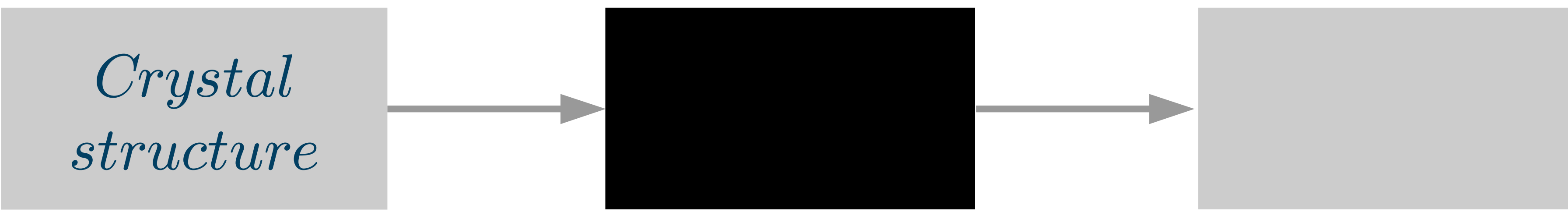
Merge simulations, physical laws, and experimental data

Reduce the need for expensive experimental development

Accelerate materials and drugs discovery

Generic with **proven** applications in materials discovery and drug design

Neural network: train on complete data



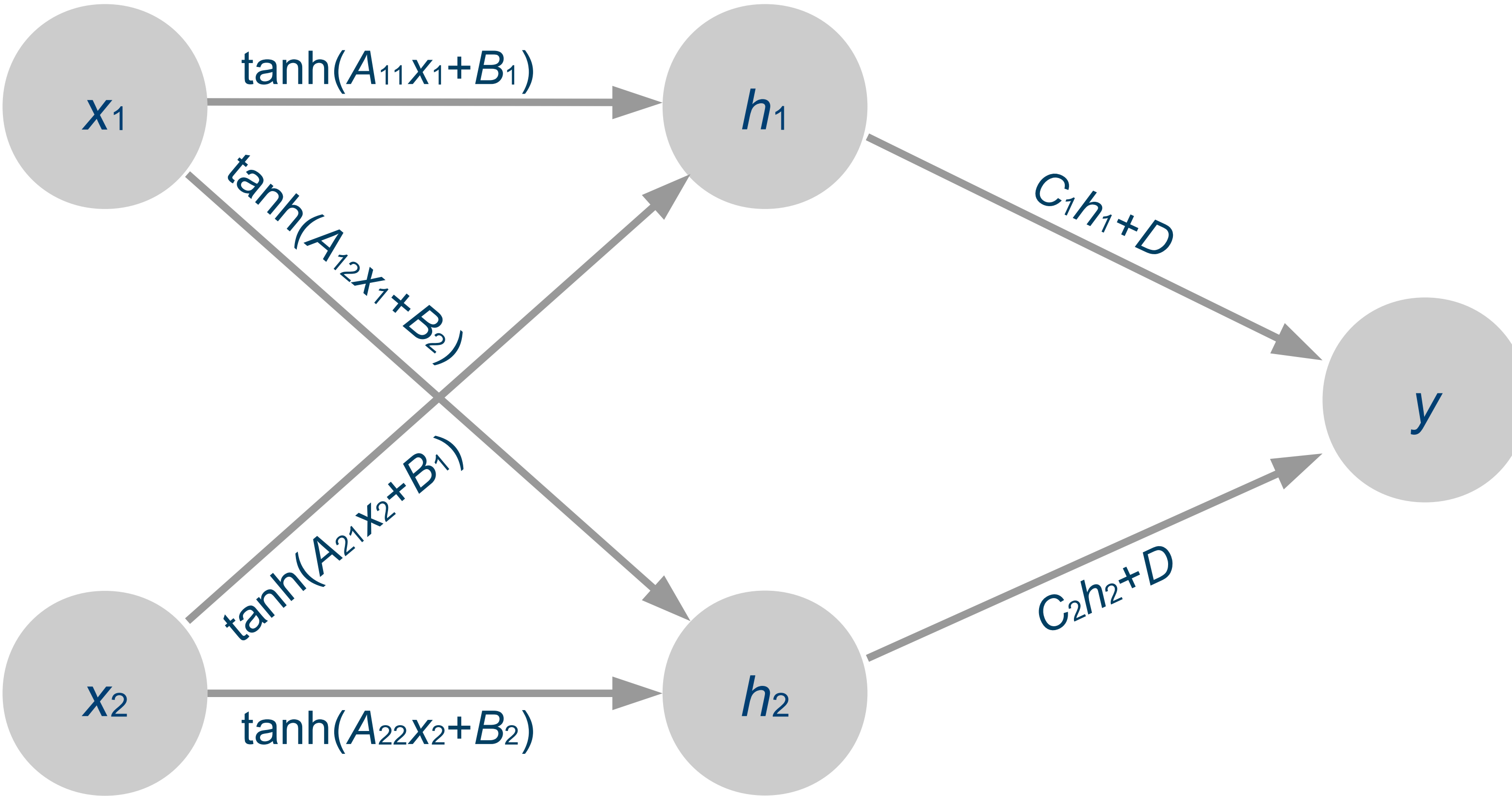
Neural network: train on complete data

Materials design

`Materials design`

Neural network: train on complete data

Neural networks: mathematical form

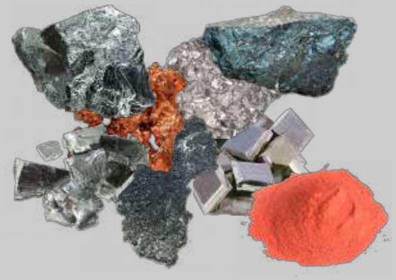


Proposed neural network: train on fragmented data

Proposed neural network: predict on fragmented data

Black box for materials design

Composition



Properties

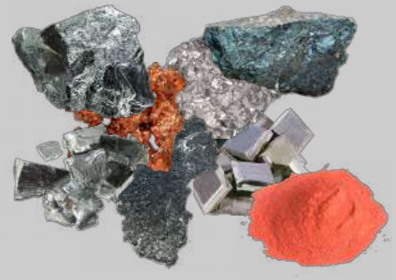
UTS

Hardness

Cost

Training the neural network

Composition



Properties

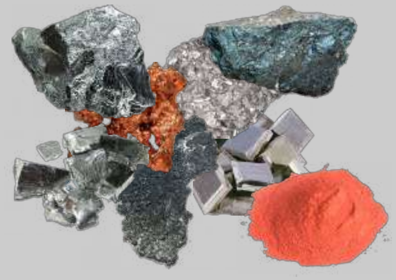
UTS

Hardness

Cost

Neural network for materials design

Composition



Properties

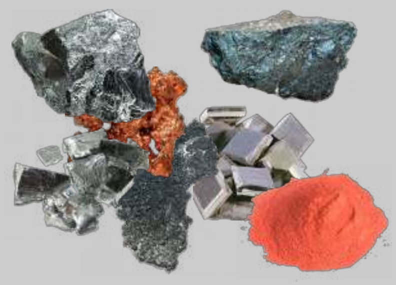
UTS

Hardness

Cost

Neural network to exploit all available correlations

Composition



Properties

UTS

Hardness

Cost

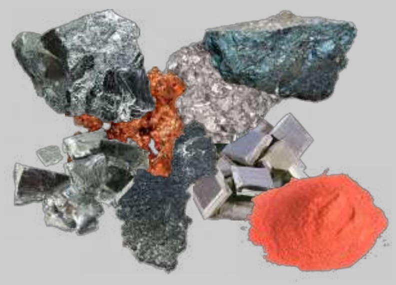
Simulations

Density Functional

Molecular dynamics

Finite element

Composition



Properties

UTS

Hardness

Cost

Simulations

Density Functional

Molecular dynamics

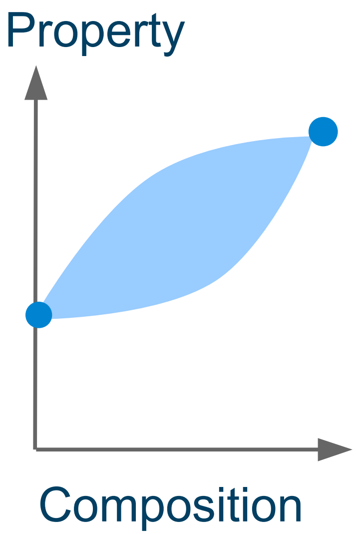
Finite element

Fragmented training data set

Composition	UTS	Hardness
✔	✔	✔
✔	✘	✔
✘	✔	✔
✔	✘	✘

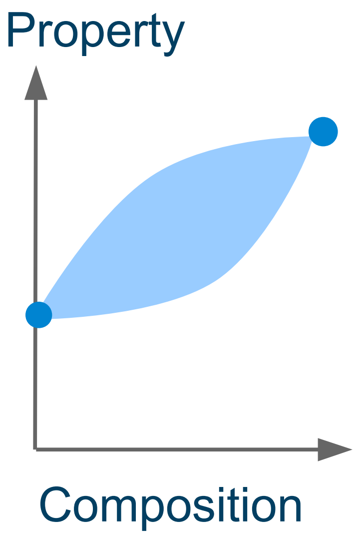
Neural network trained on experimental data

Experiment

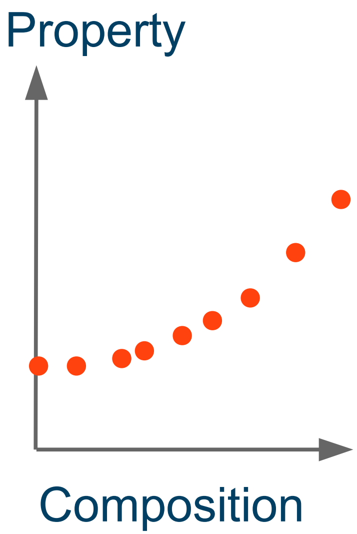


Further information is provided by a simulation

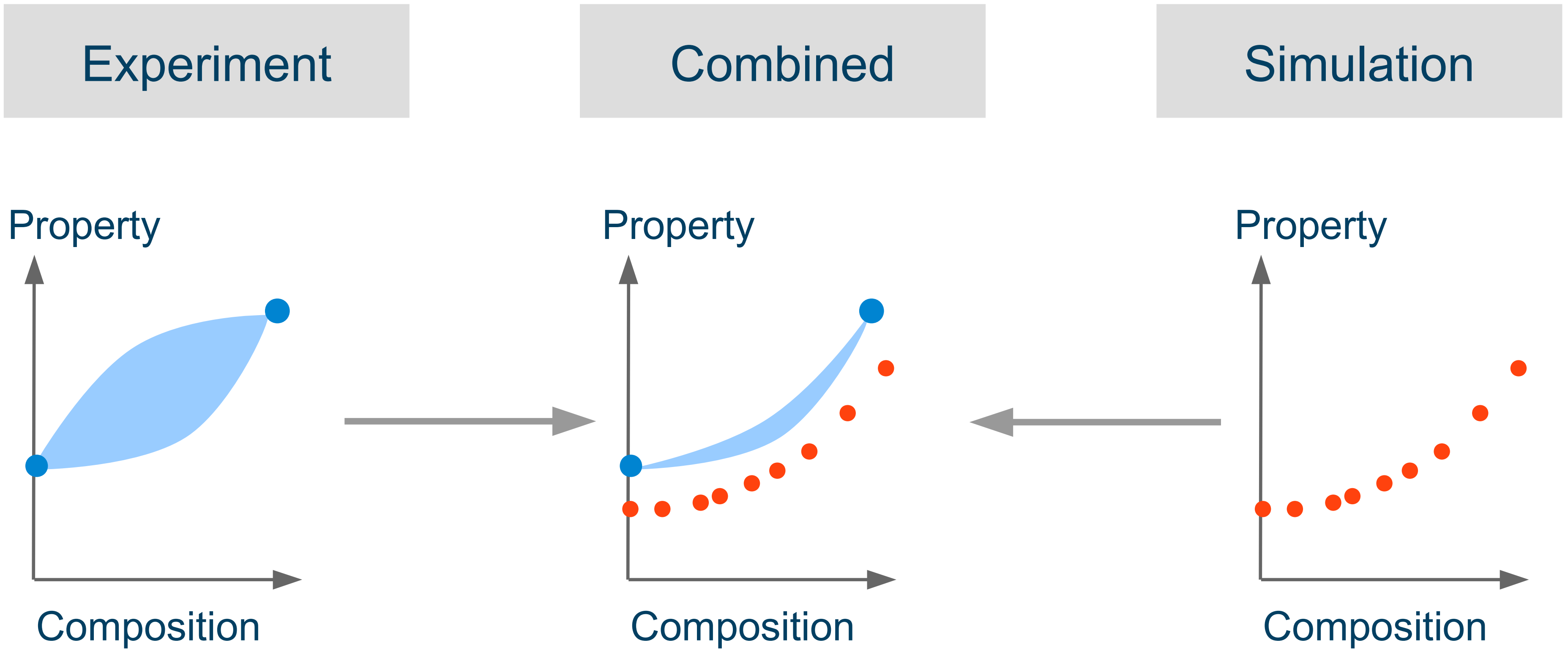
Experiment



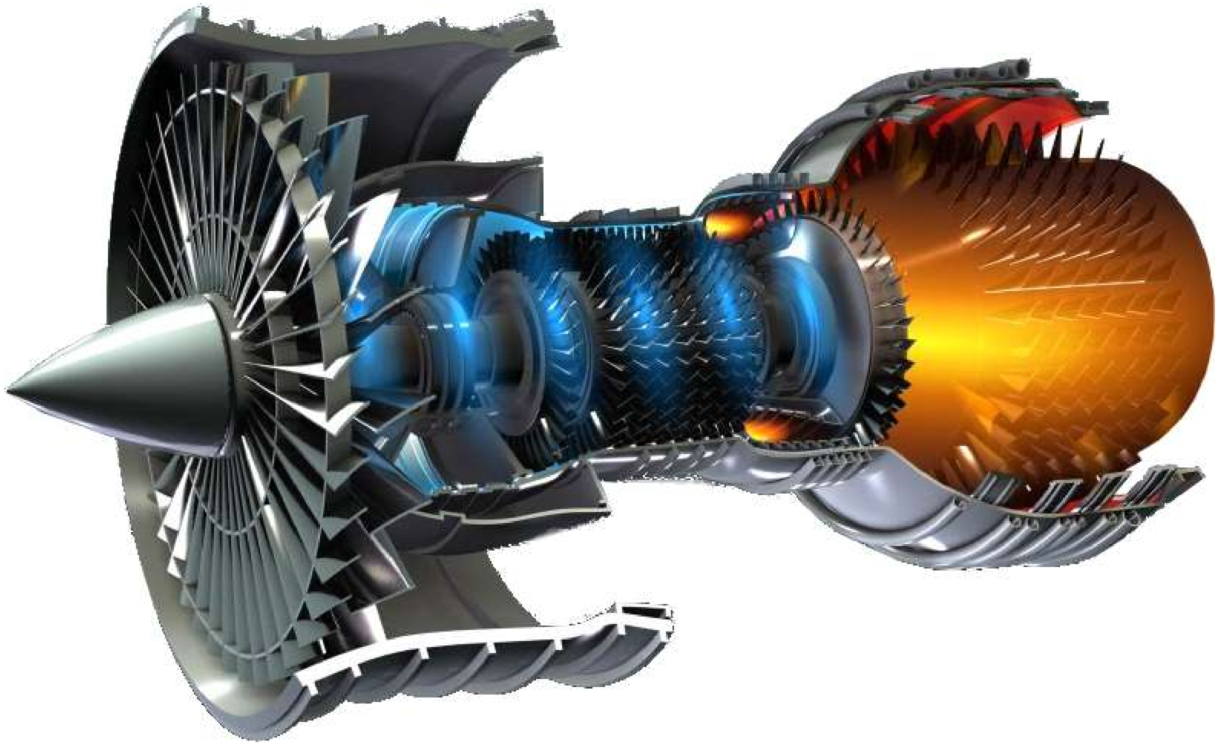
Simulation



Neural network combines the two sources of data



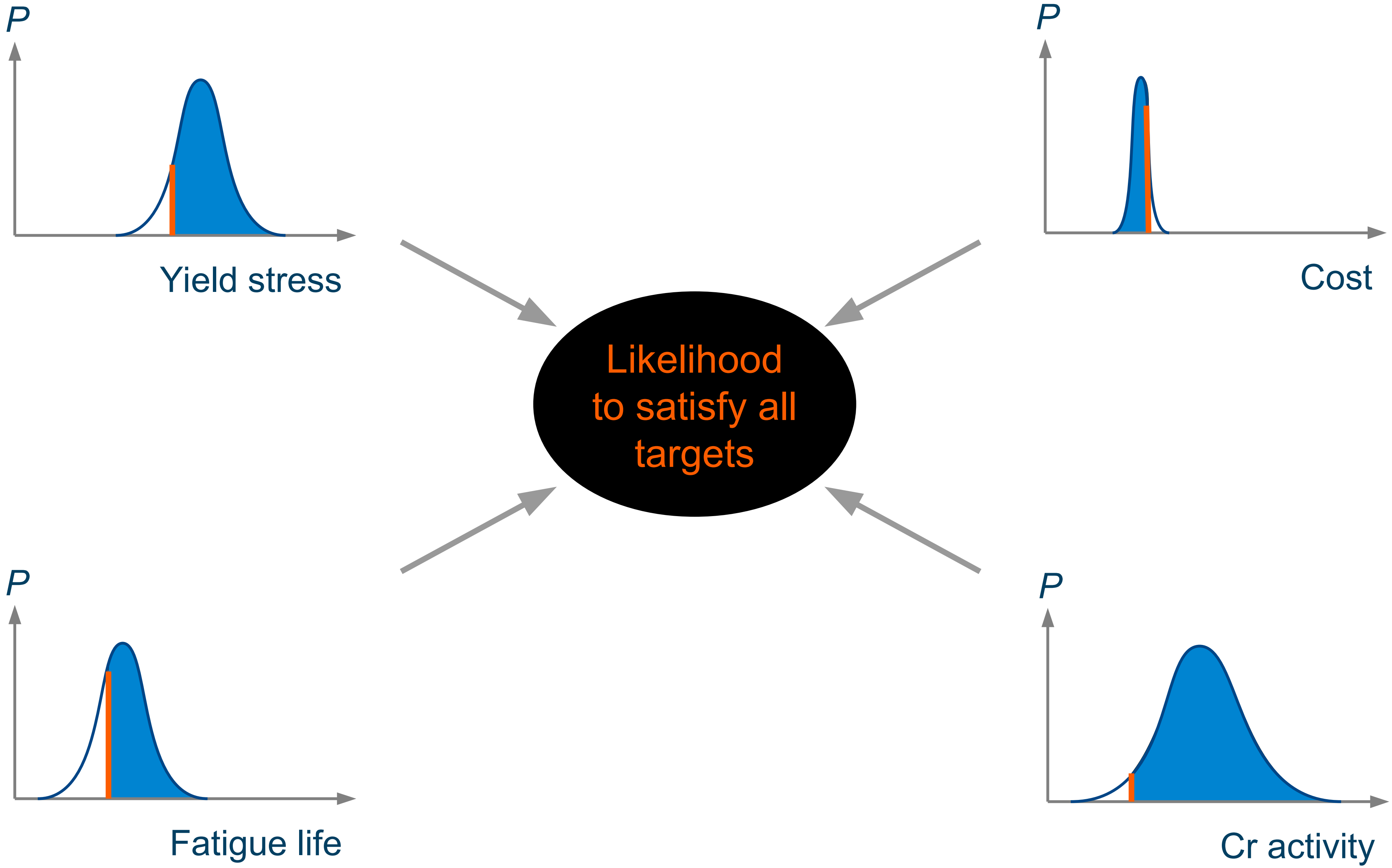
Schematic of an engine



Target properties

Cost	< 33.7 $\$kg^{-1}$
Density	< 8281 kgm^{-3}
γ' content	< 50.4 vol%
Phase stability	> 99.0 vol%
Fatigue life	> $10^{3.9}$ cycles
Yield stress	> 752.2 MPa
Ultimate tensile strength	> 960.0 MPa
300hr stress rupture	> 674.5 MPa
Cr activity	> 0.14
γ' solvus	> 983ºC
Tensile elongation	> 11.6%

Maximize the likelihood of success



Proposed alloy

Cr:15.8

Co: 20.0

Mo: 0.5

W: 0.5

Ta: 4.9

Nb: 1.1

Al: 2.4

Ti: 3.0

Fe: 3.9

Mn: 0.2

Si: 0.2

C: 0.02

B: 0.06

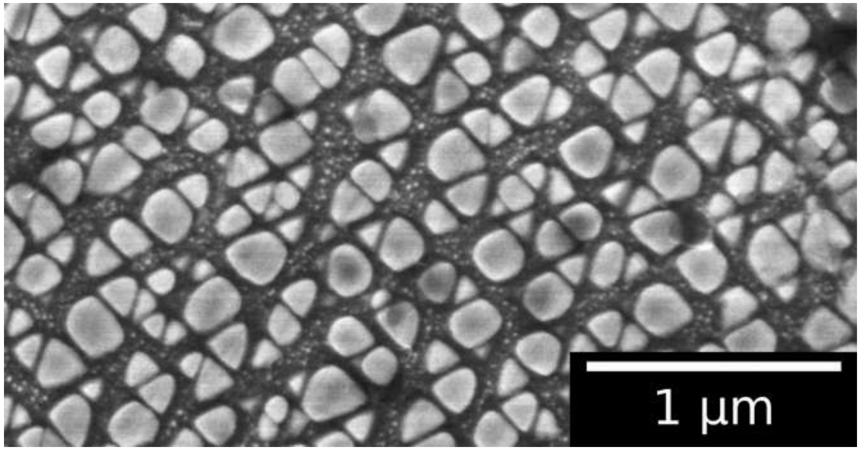
Zr: 0.18

Ni: 47.2

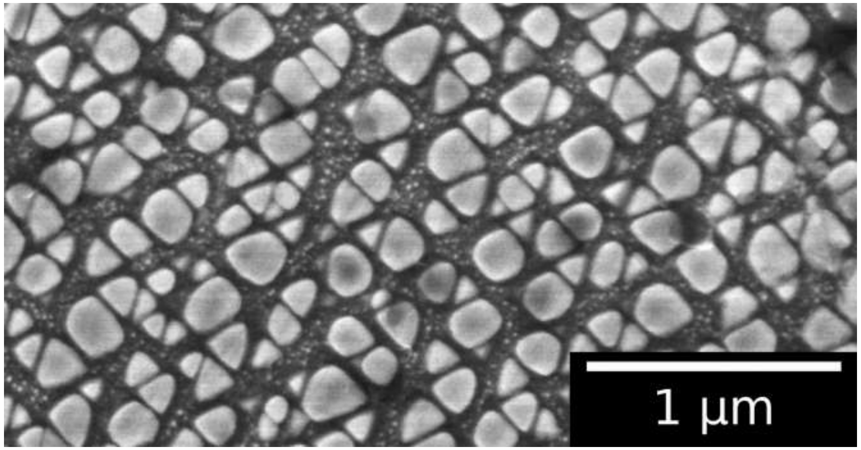
900ºC

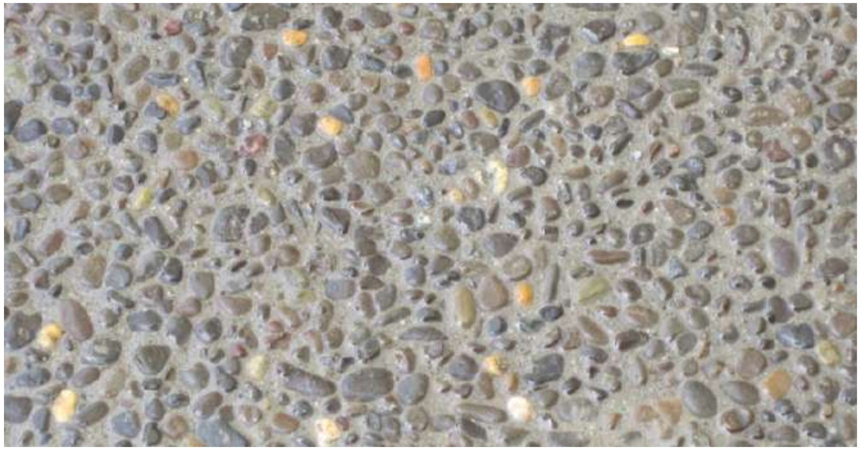
30 hours

Microstructure

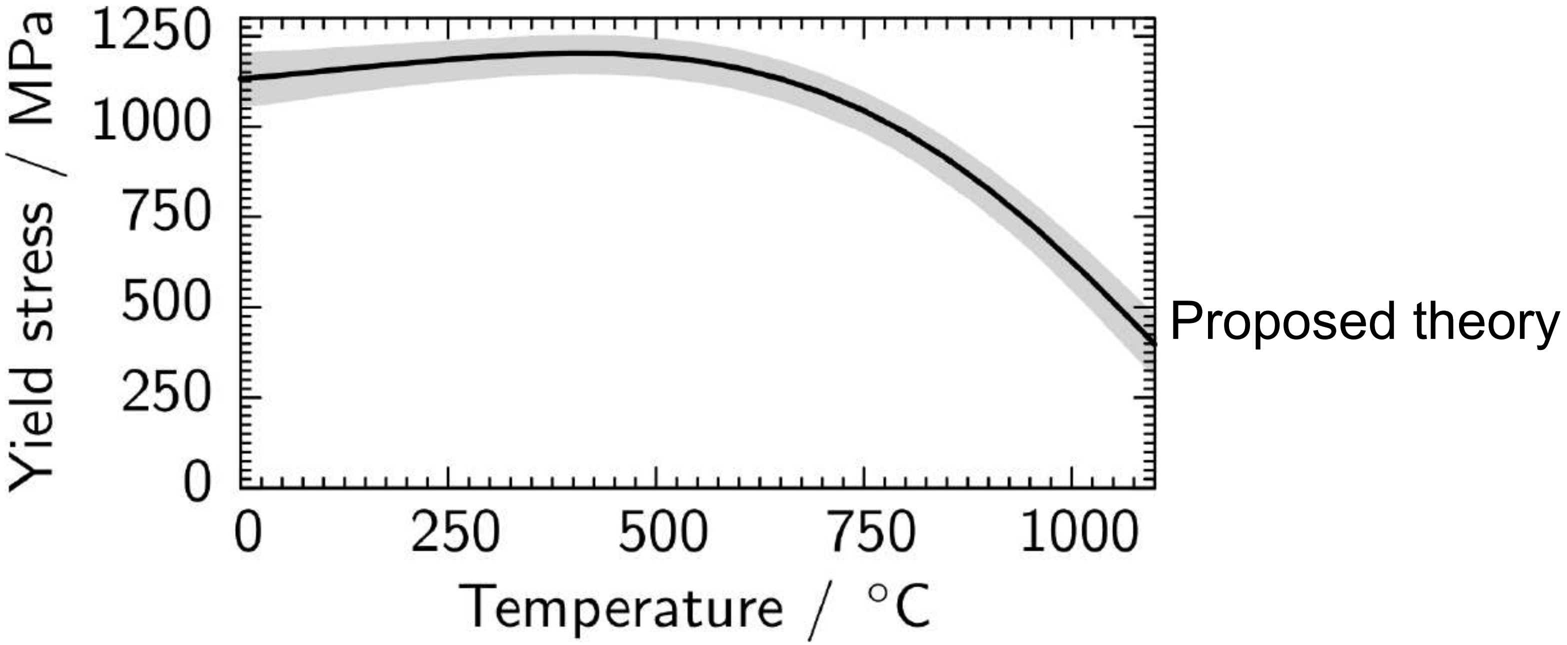


Precipitates strengthen the alloy

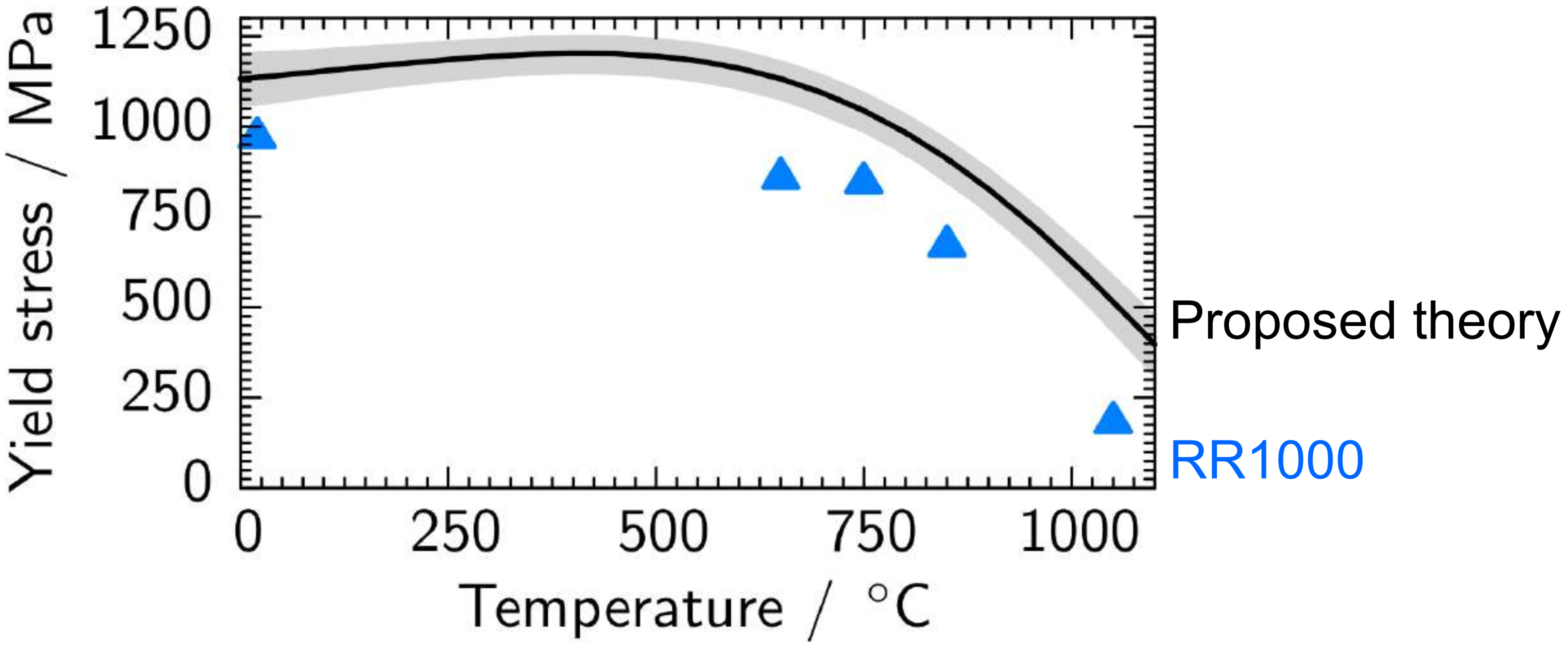




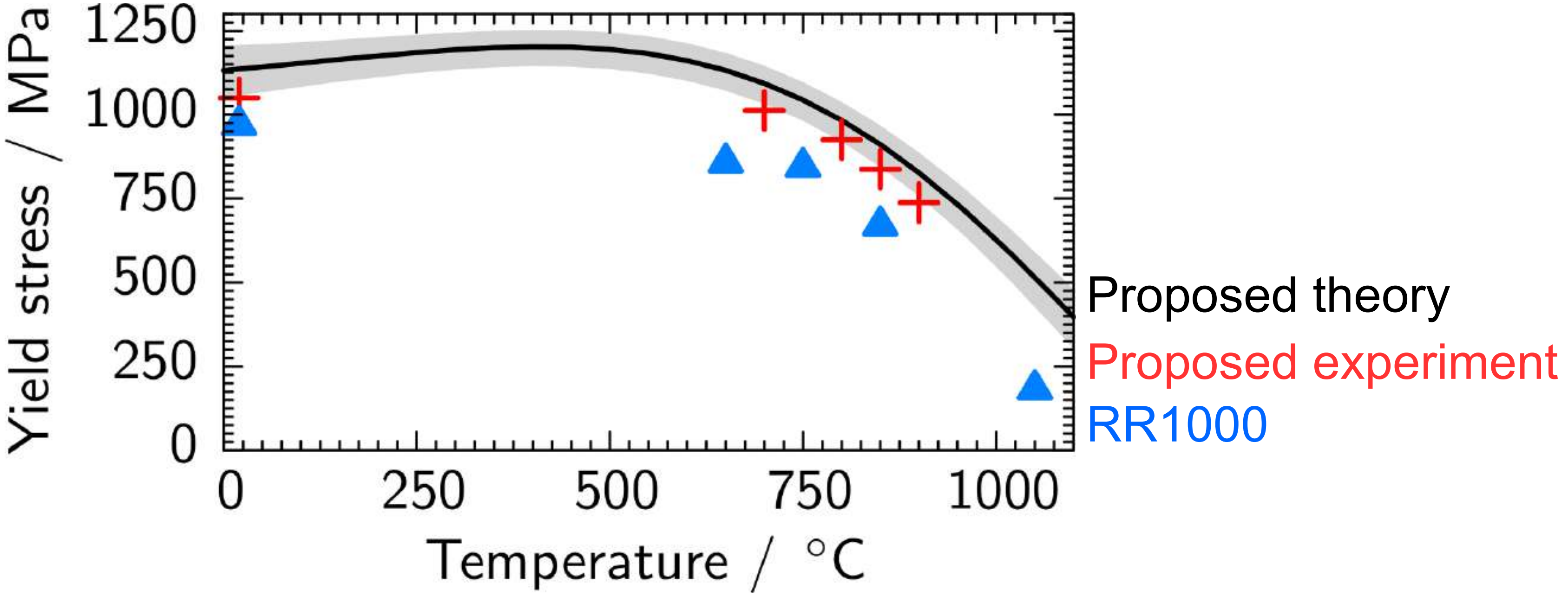
Predict the yield stress



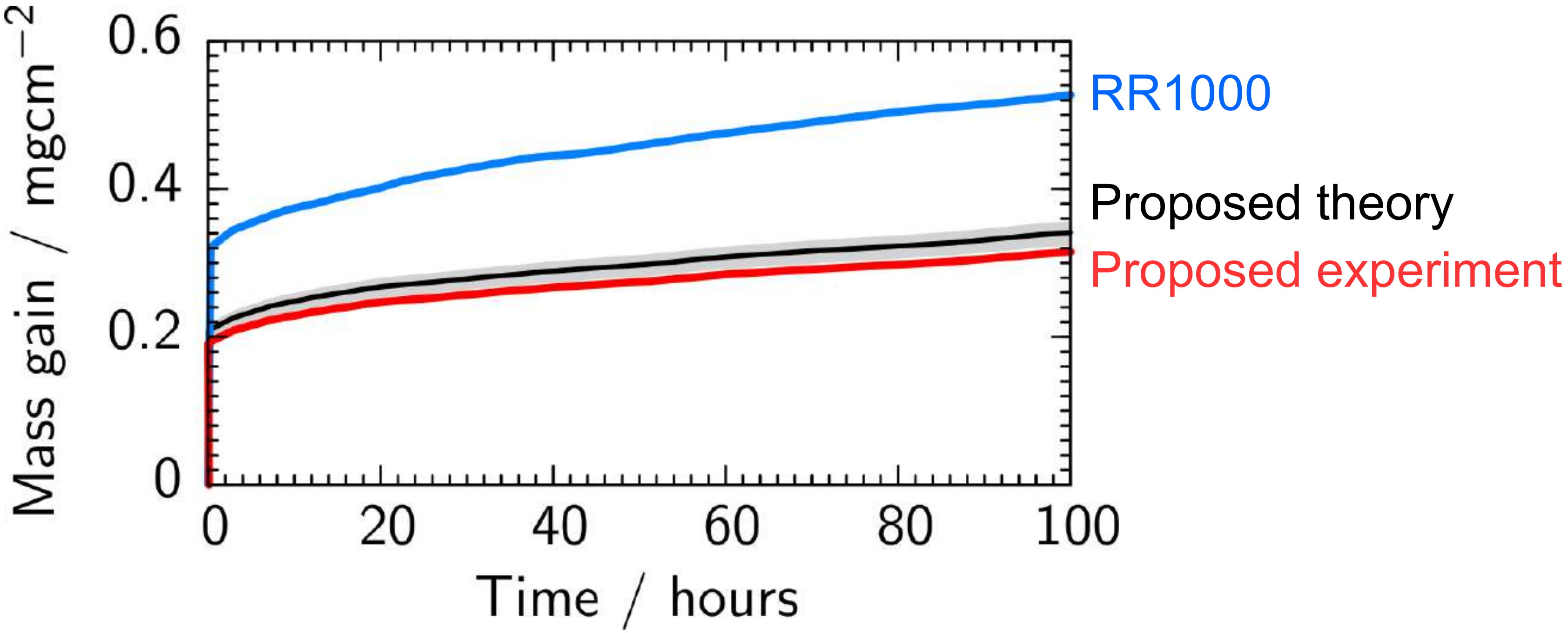
Test the yield stress



Test the yield stress

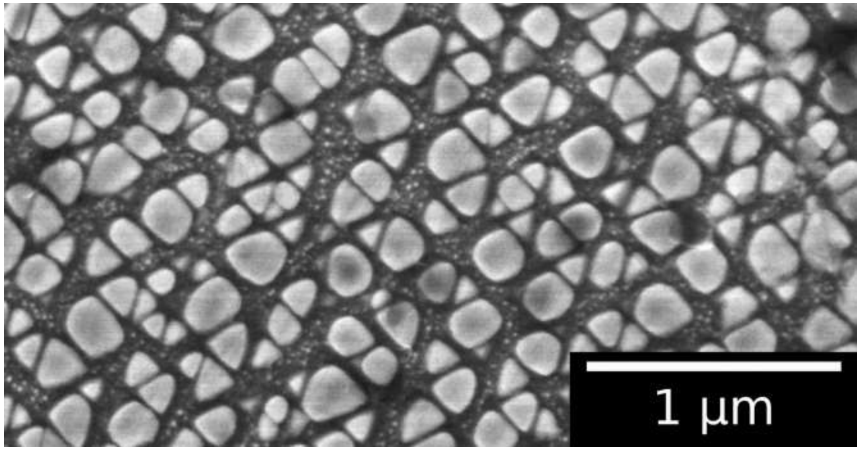


Test the oxidation resistance

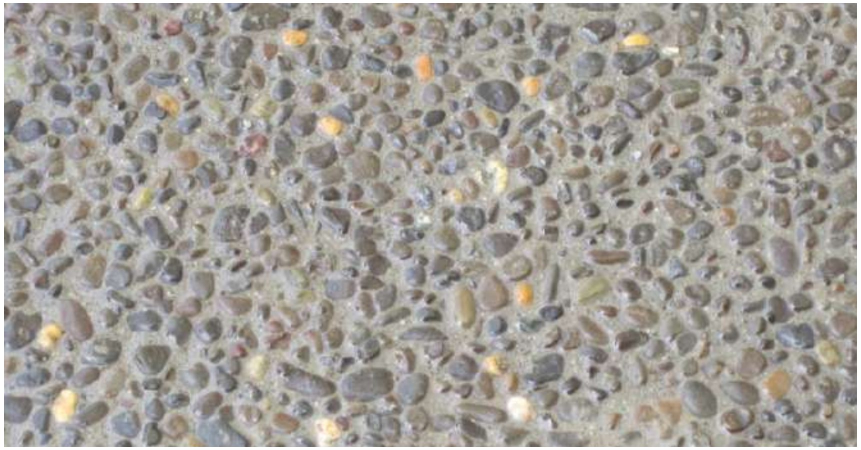


Microstructure strengthens the alloy

Precipitates

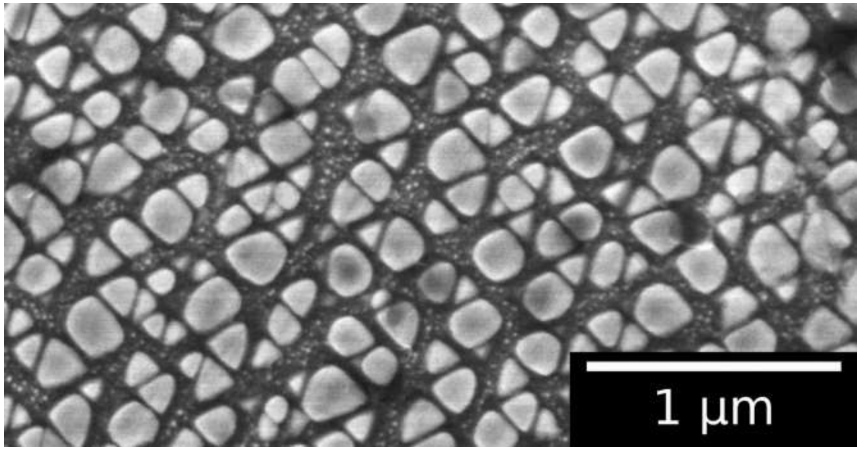


Aggregate

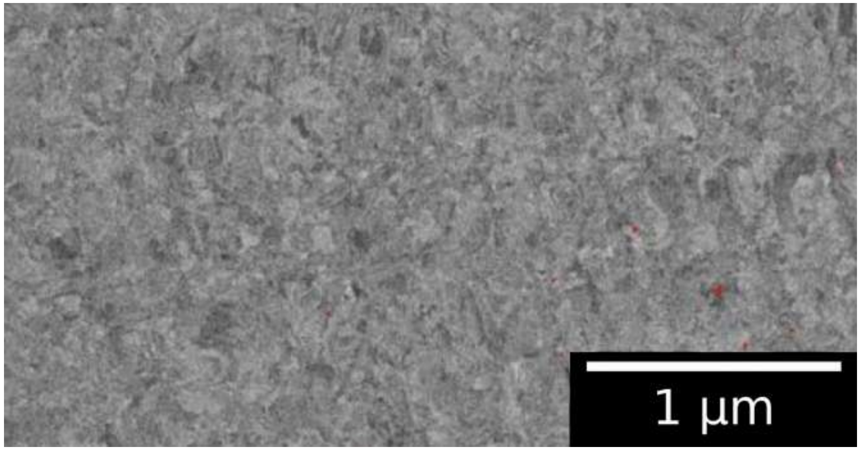


Microstructure defined by the heat treatment

Correct heat treatment

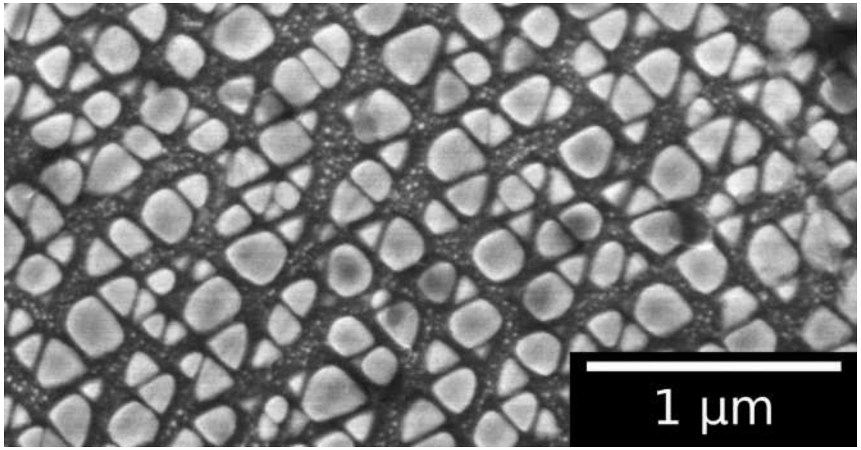


Incorrect heat treatment

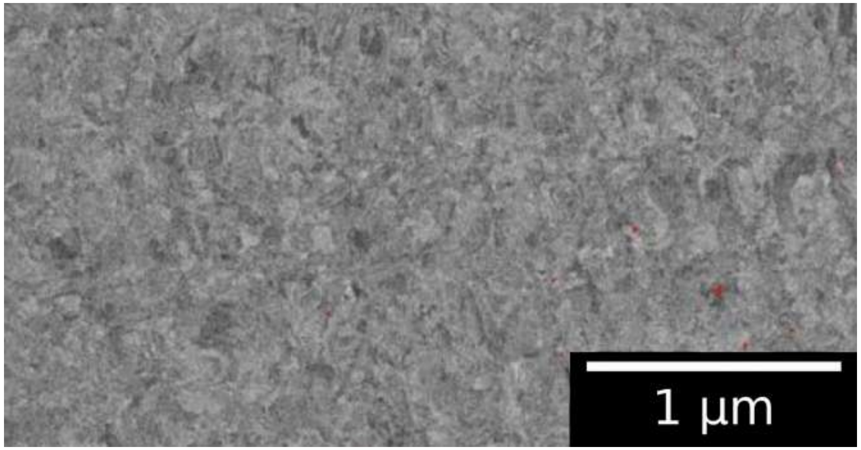


Microstructure distribution links to the heat treatment

Correct heat treatment

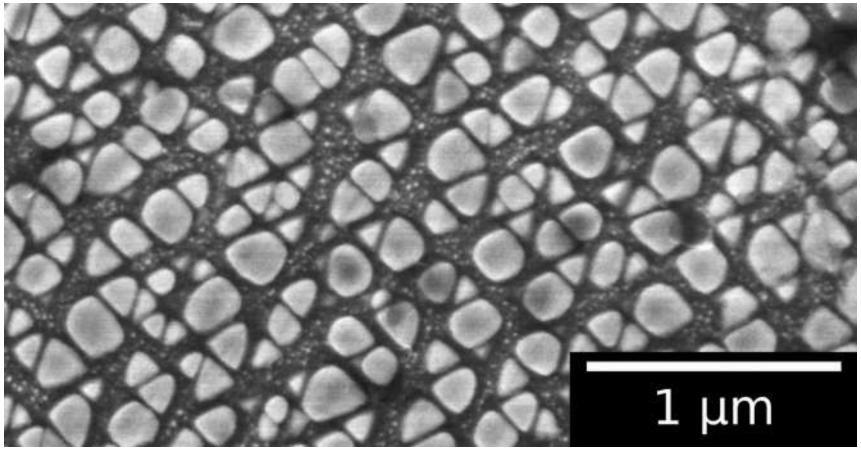


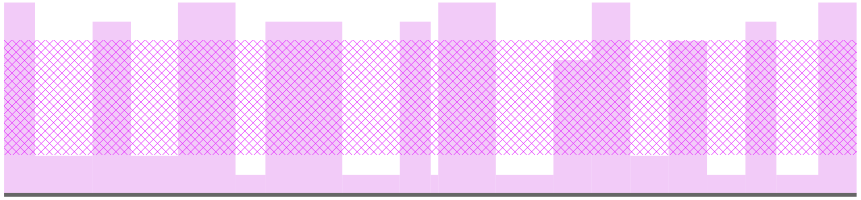
Incorrect heat treatment



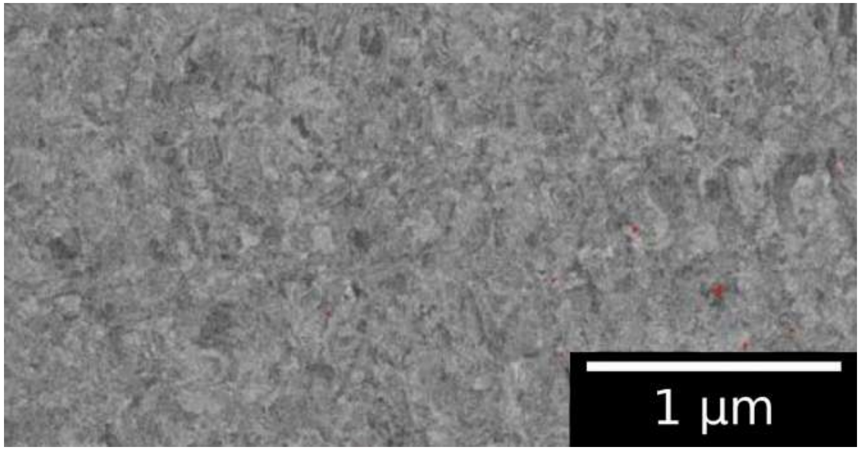
Microstructure distribution links to the heat treatment

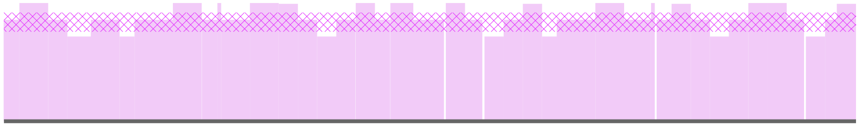
Correct heat treatment



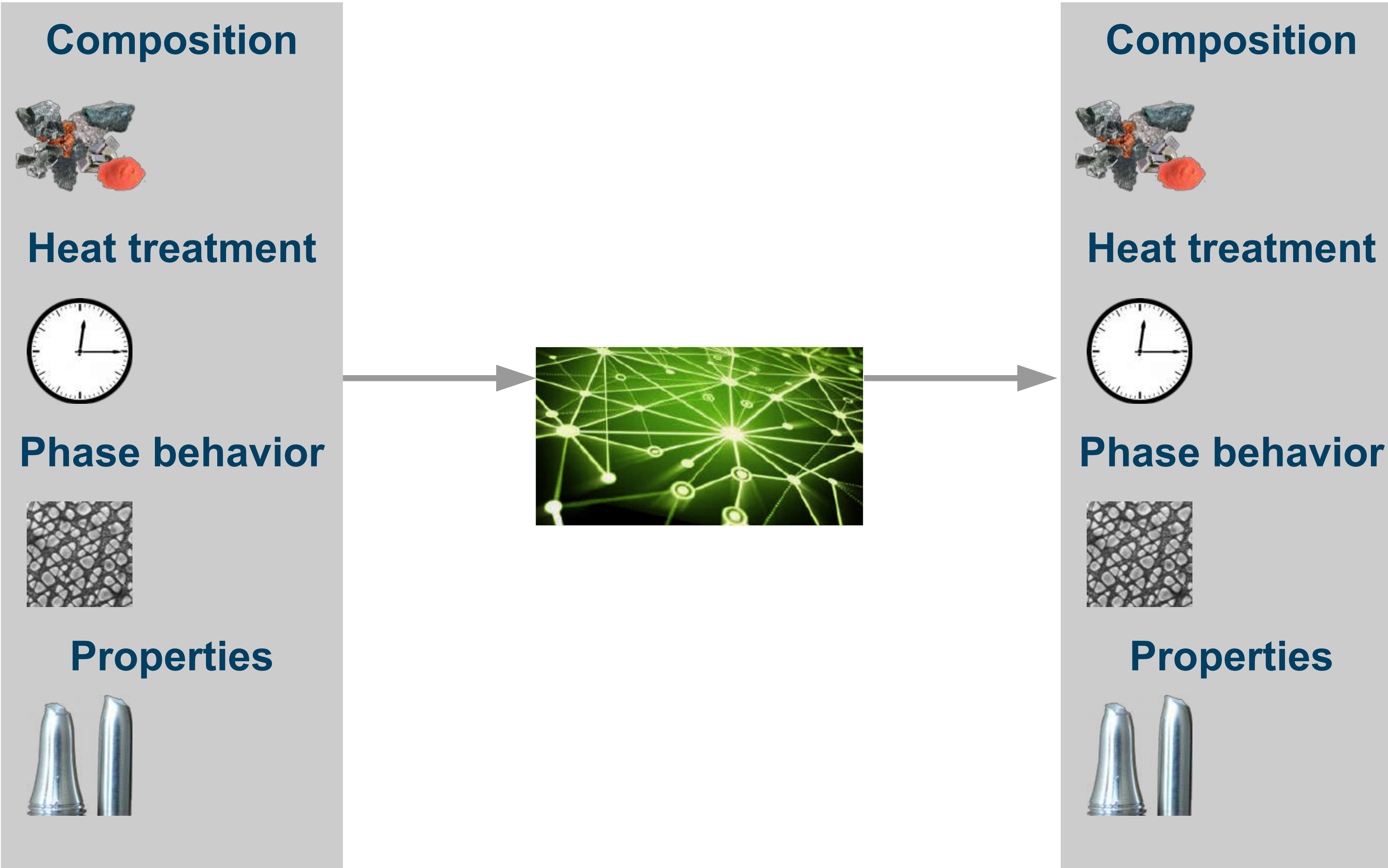


Incorrect heat treatment

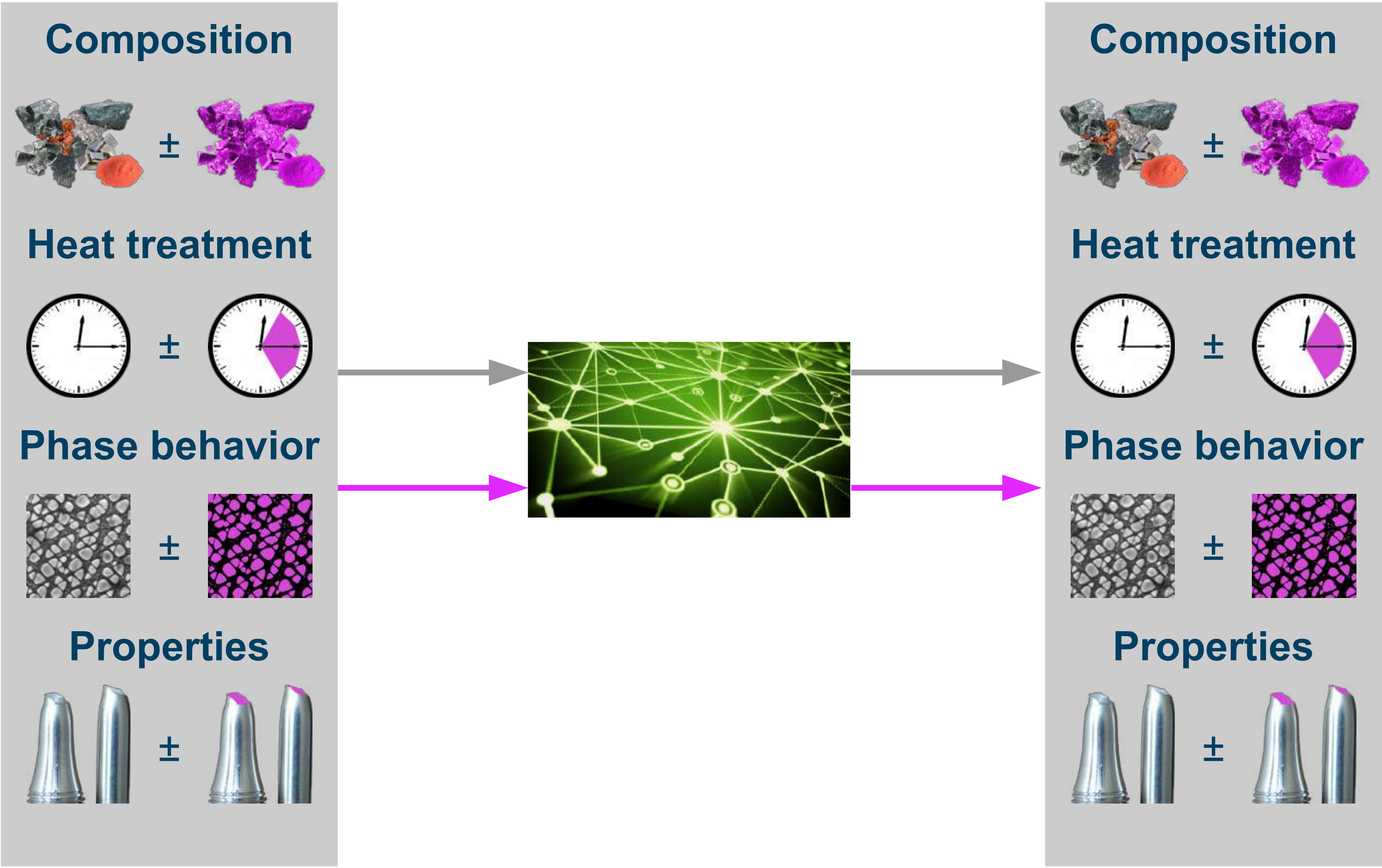




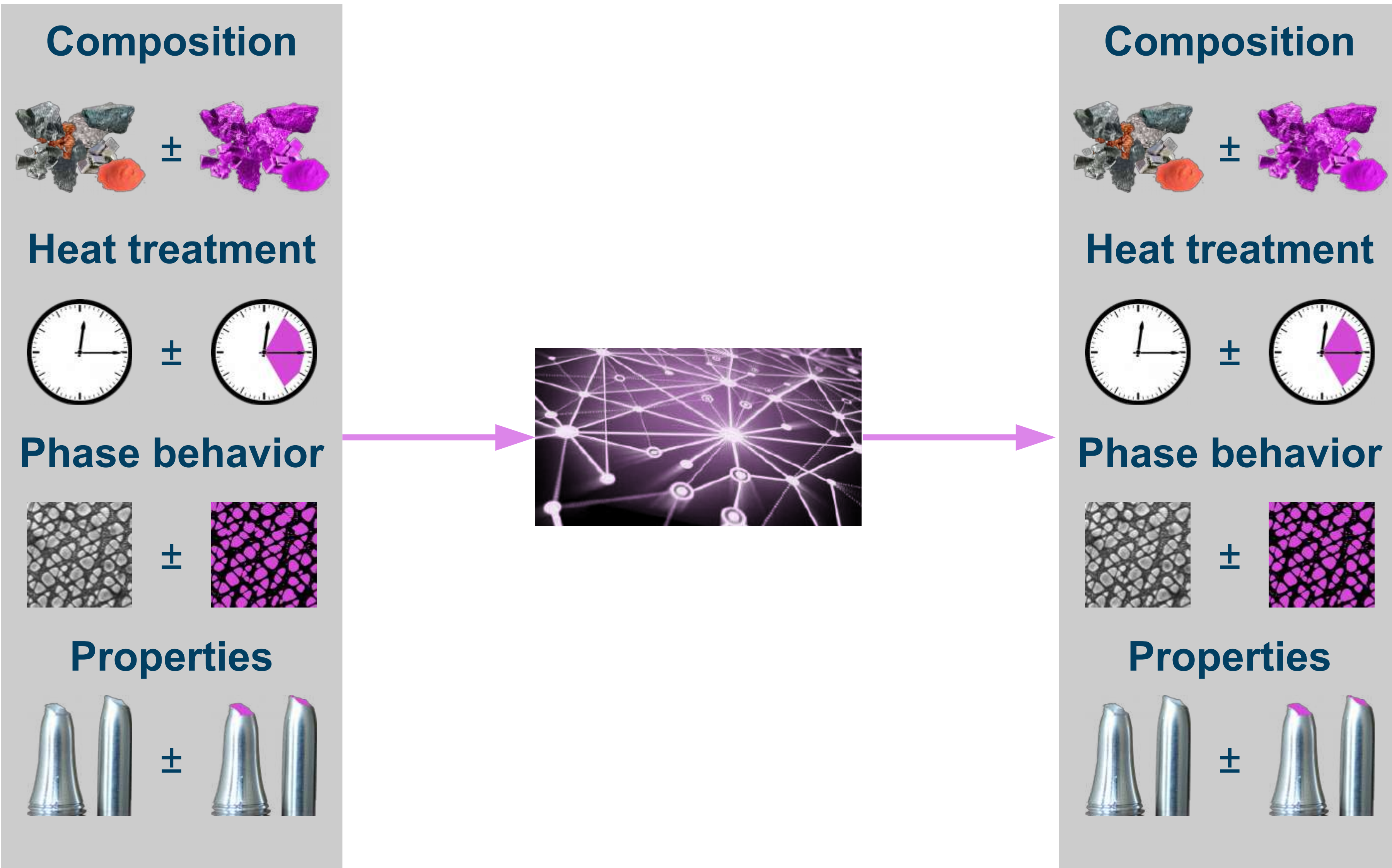
Standard neural network



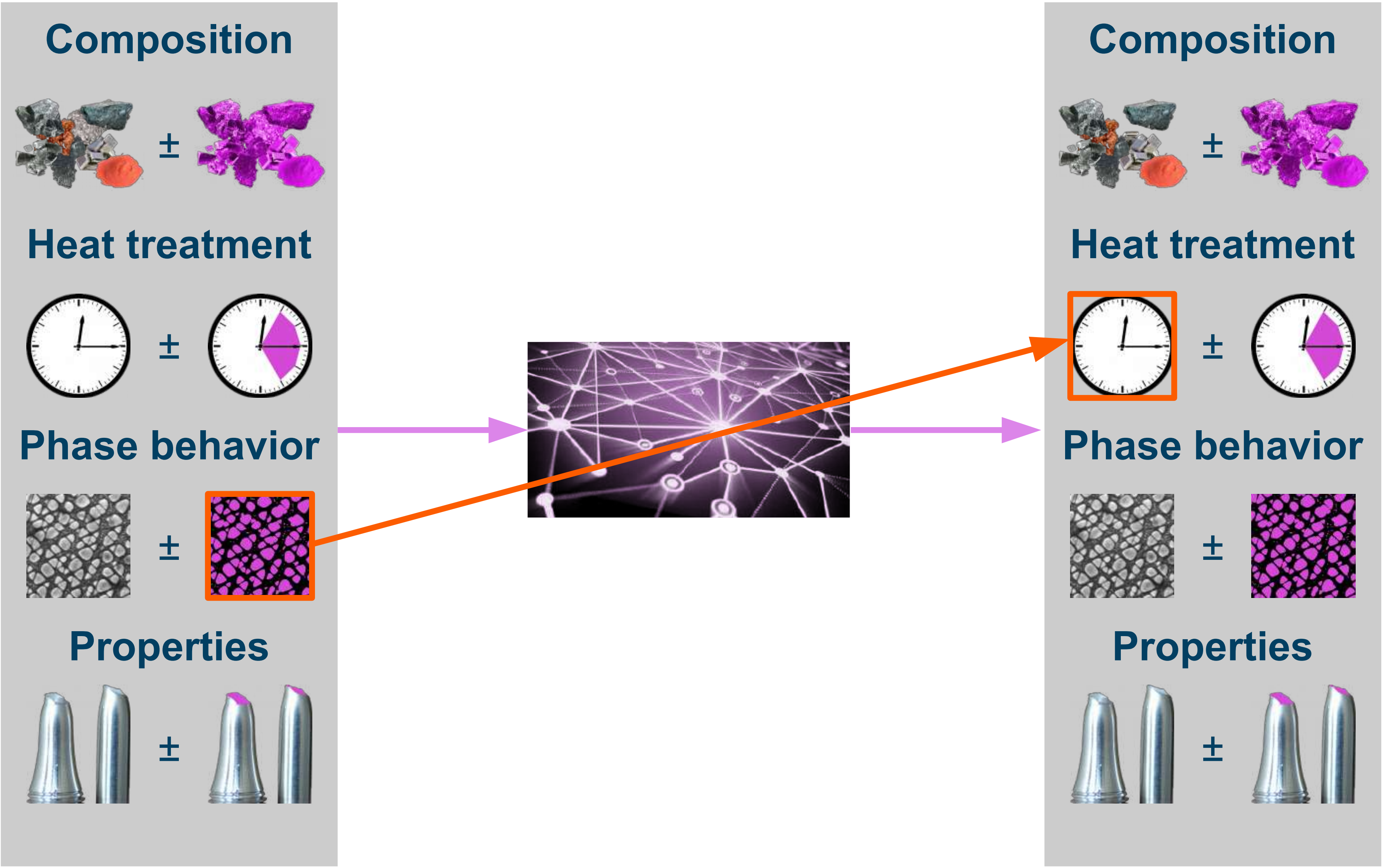
Neural network transmits noise as uncertainty



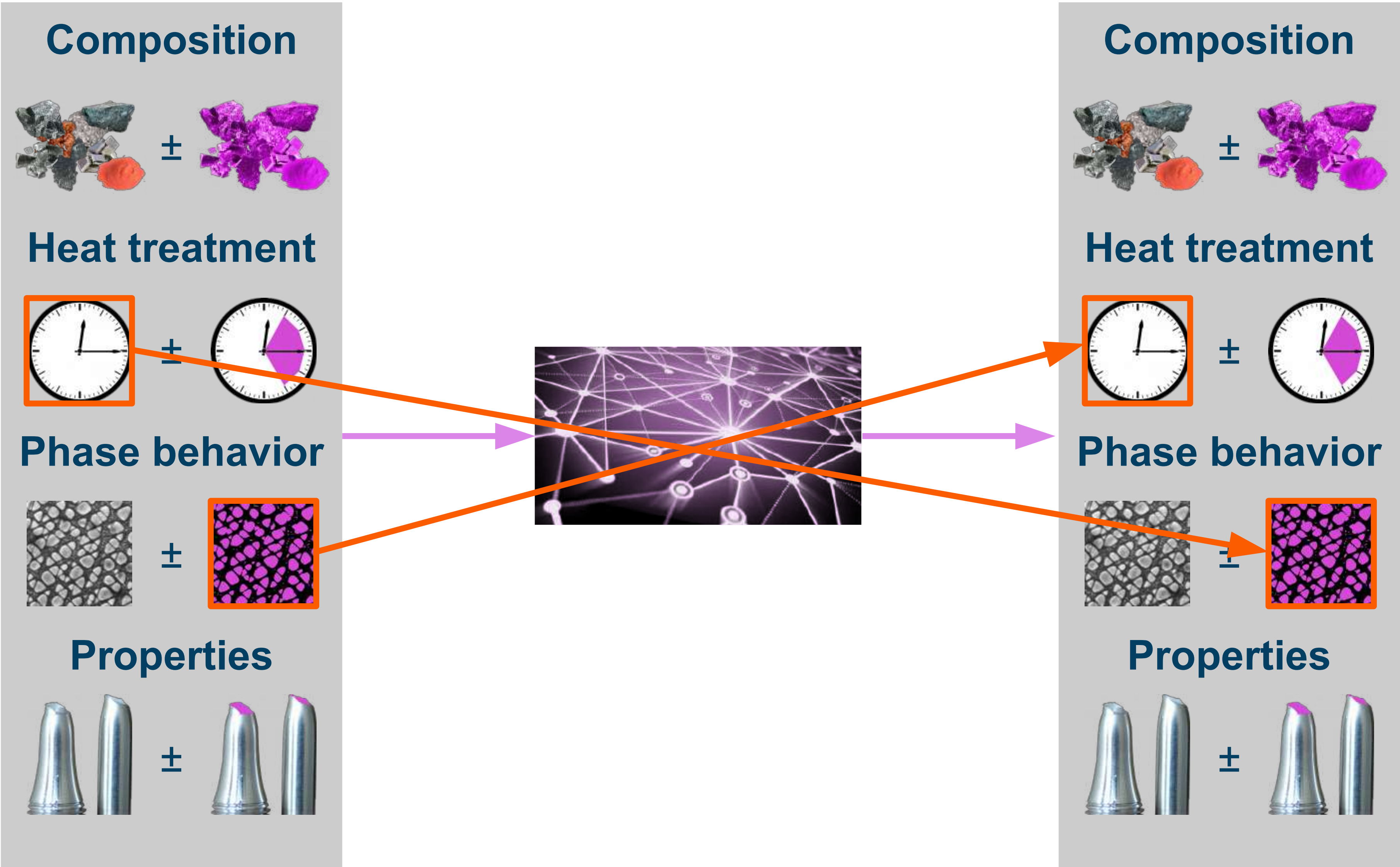
Incorporate noise into the neural network



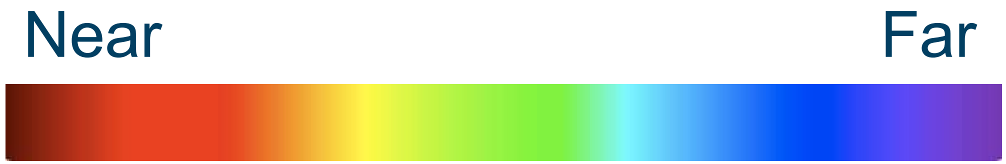
Exploit noise in the neural network



Exploit noise in the neural network



Point cloud: noise in the data



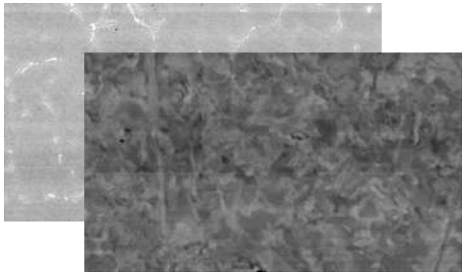
Point cloud: benefits of including noise

Recover presence of people to 90% accuracy

Also applies to trees and railings

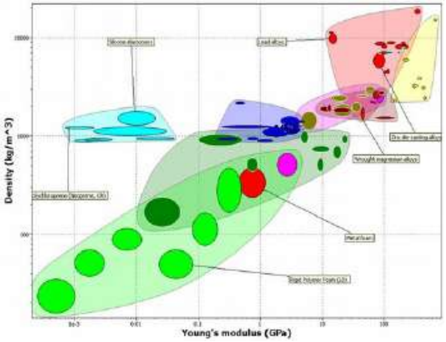
More materials designed

Molybdenum forging alloys



3D printed alloy designed from 7 data entries

Found 192 errors in materials databases

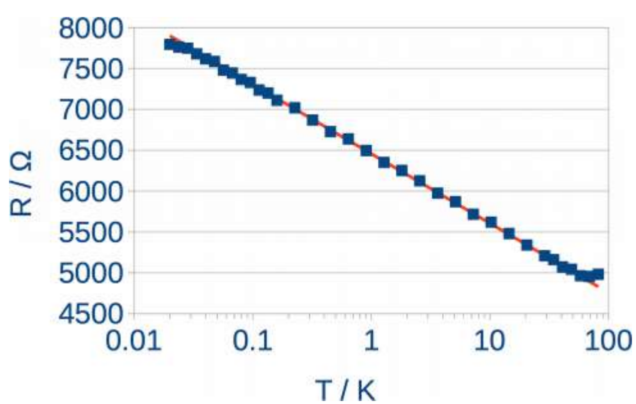


Even more materials designed

Battery design with DFT and experimental data

Designing lubricants with DFT and experimental data

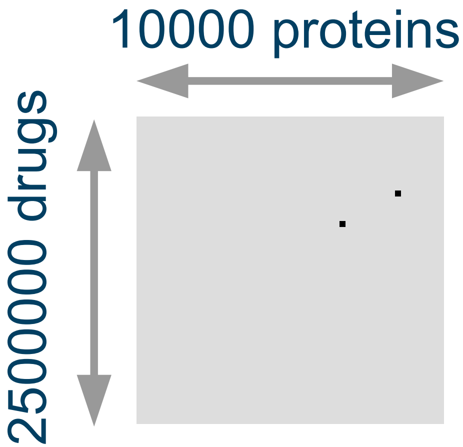
Thermometer with quantum and experimental data



Data available for drug discovery

10,000 proteins with 2,500,000 compounds

Original dataset 0.05% complete



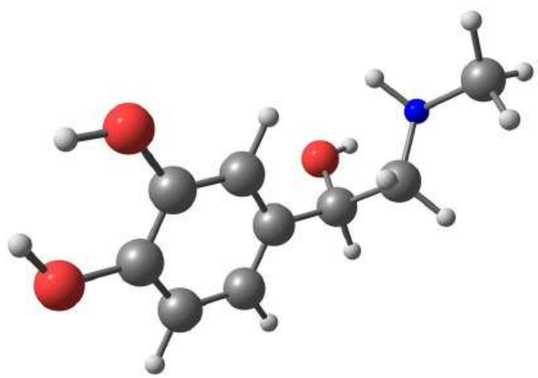
Impute the database used for drug discovery

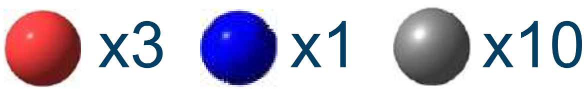
10,000 proteins with 2,500,000 compounds

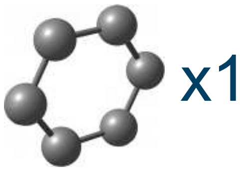
Original dataset 0.05% complete

Filled 32% of the entries

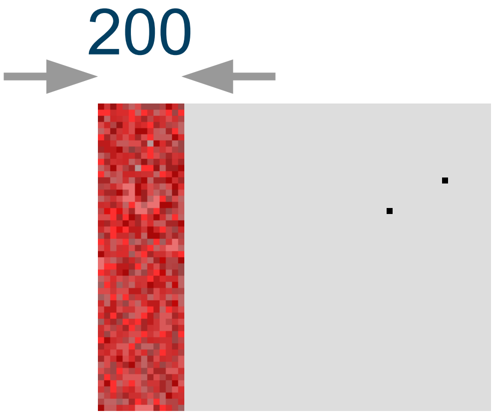
Drug discovery with additional descriptors







1101 Morse descriptor

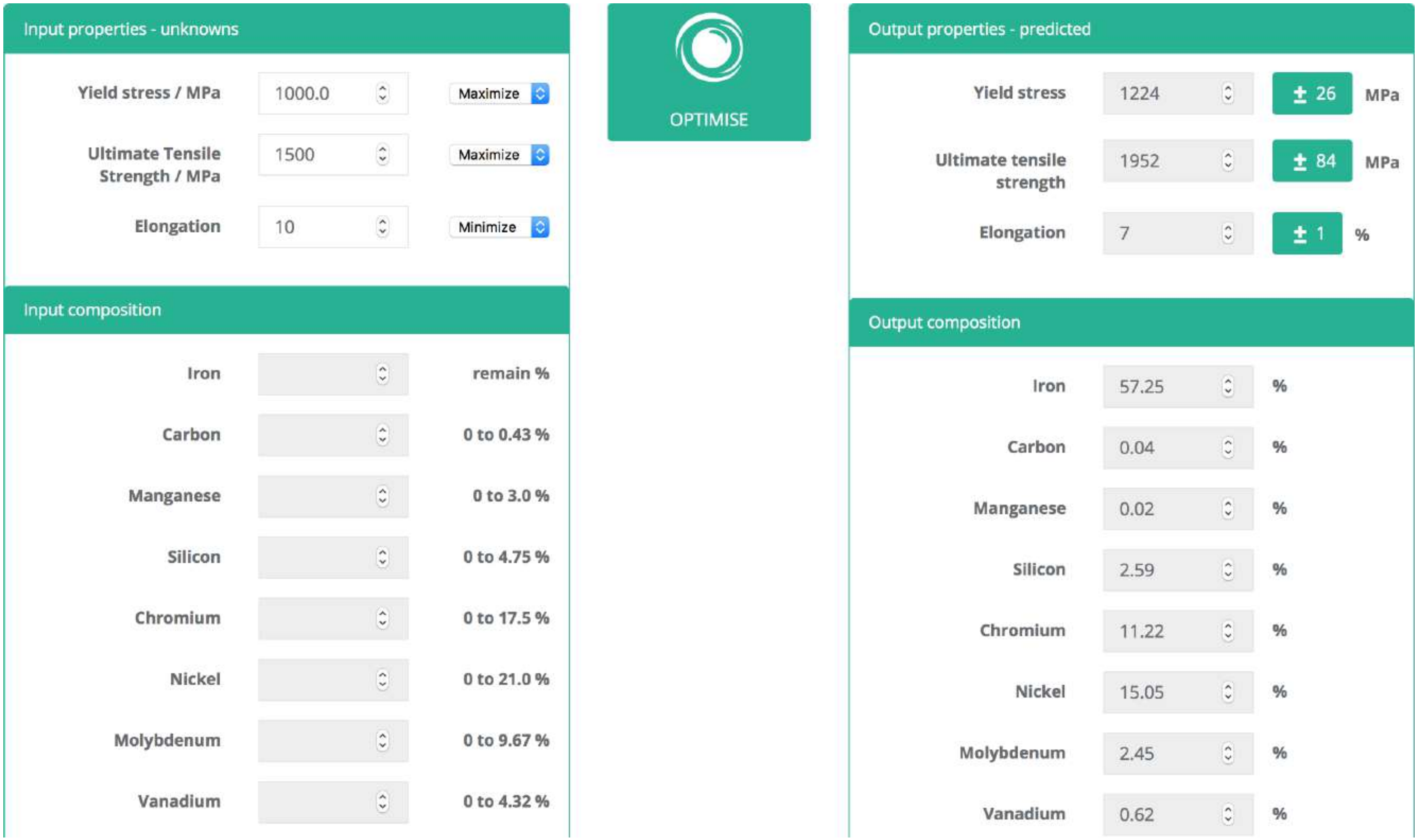


Improved drug discovery

Include drug structural information to fill to 46%

Saved >$1billion in experimental costs

Startup intellegens productizing the neural network



Summary: progress

Apply deep learning to high-value **fragmented** data

Exploit knowledge of **probability distribution** of the data

Experimentally **proven** materials and drugs design with 7 companies, founded startup **intellegens**

Summary: opportunities

Apply deep learning to high-value fragmented data

Exploit knowledge of probability distribution of the data

Experimentally proven materials and drugs design with 7 companies, founded startup intellegens

Merge experiments and simulations into holistic design tool

Summary: future prospects

Apply deep learning to high-value fragmented data

Exploit knowledge of probability distribution of the data

Experimentally proven materials and drugs design with 7 companies, founded startup intellegens

Merge experiments and simulations into holistic design tool

Engineers establish all possible sources of information