UNIVERSITY OF CAMBRIDGE

Ontologies for machine learning driven materials design

Gareth Conduit

Machine learning can

Model datasets where relationships are complex and no analytical model exists

Exploit property-property relationships

Merge computer simulations and physical laws with experimental data

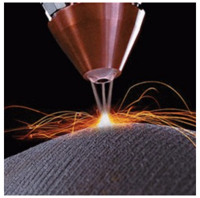
Reduce costly experiments to accelerate discovery

Black box machine learning for materials design

Composition

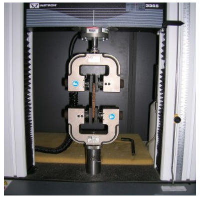
Properties

Defects

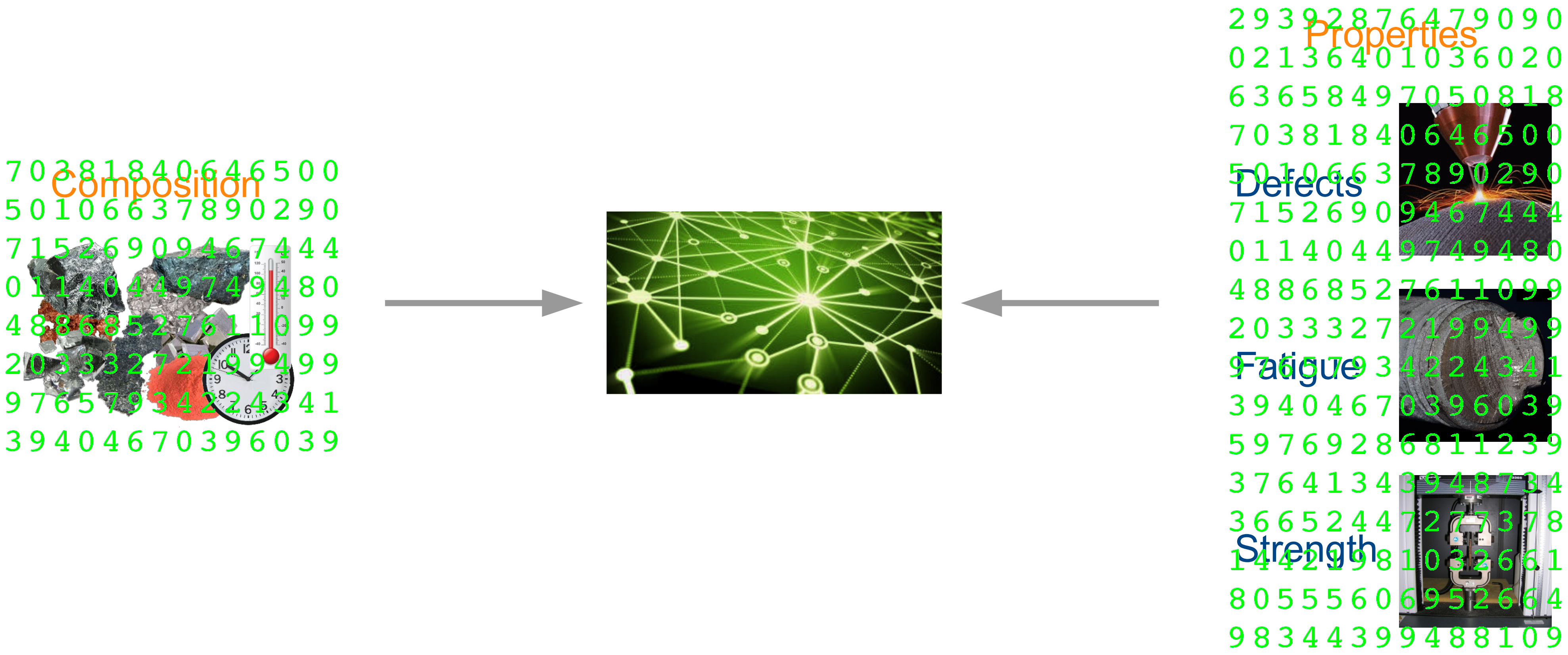


Fatigue

Strength



Train the machine learning

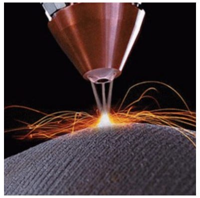


Machine learning predicts material properties

Composition

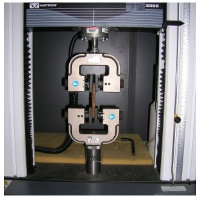
Properties

Defects



Fatigue

Strength



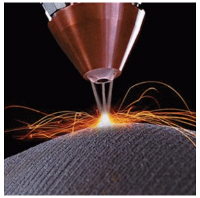
Exploit property-property relationships

Composition

Feedback
property value

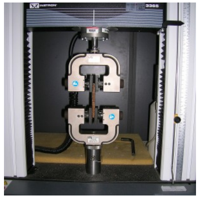
Properties

Defects

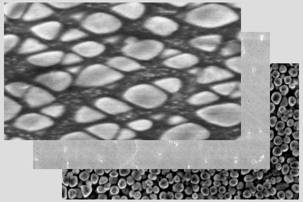


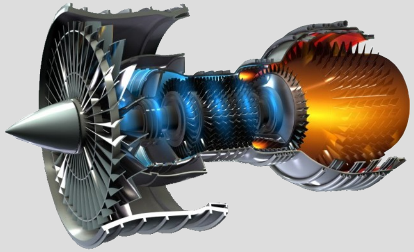
Fatigue

Strength



DOMINO



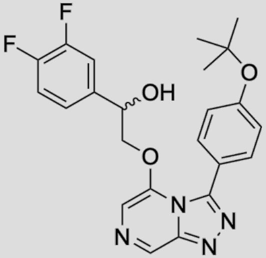


ROLLS ROYCE

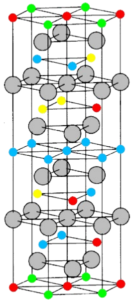
BOEING

GKN

Agency for Science, Technology and Research
SINGAPORE



e-therapeutics
optibrium
Genentech
A Member of the Roche Group
University of Dundee
OSM
Constellation PHARMACEUTICAL
zoetis
Celgene
MILLENNIUM PHARMACEUTICALS, INC
Janssen
Takeda
AstraZeneca
SANOFI



SAMSUNG

WA

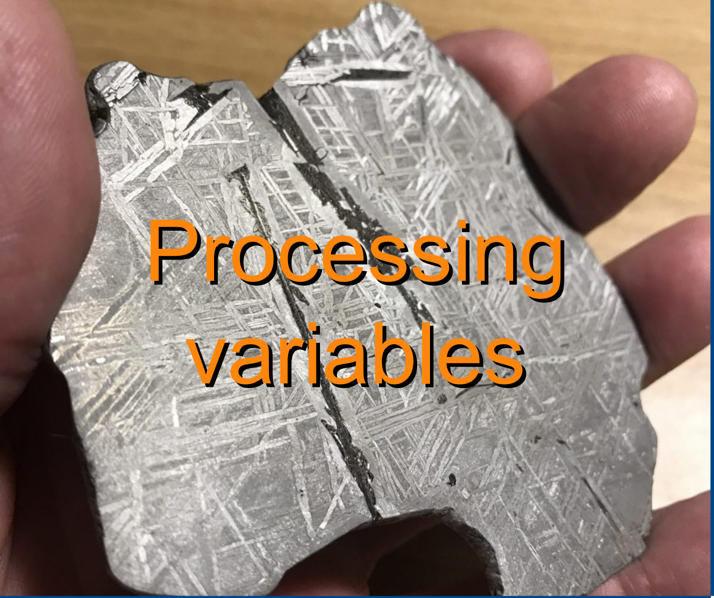
Value of a material through its development cycle

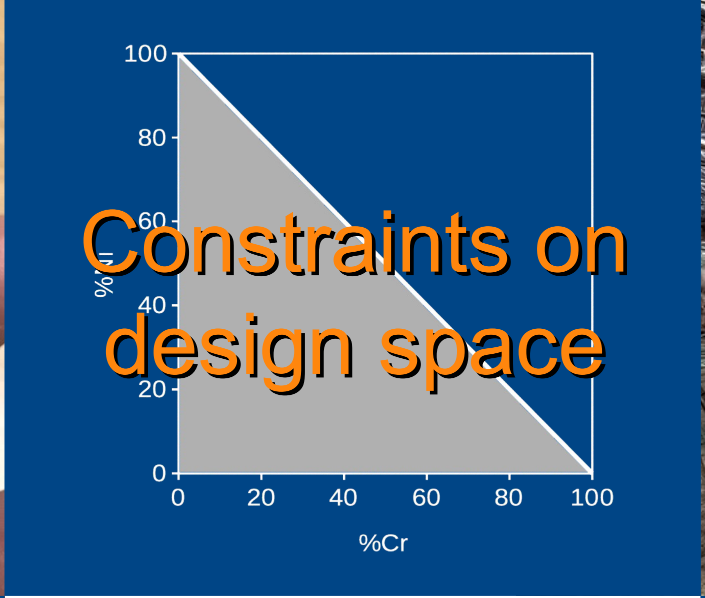
Proposal

Testing

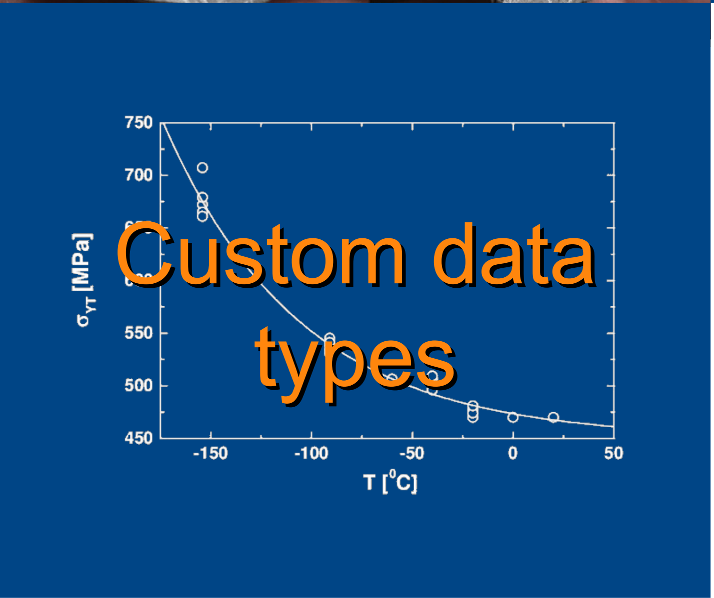
Production

Proposal	Testing	Production
10%	10%	80%


Processing variables


Constraints on design space
100
80
60
40
20
0
0 20 40 60 80 100
%Cr


Esoteric variables


Custom data types
750
700
550
500
450
-150 -100 -50 0 50
T [°C]
σYT [MPa]


Security and provenance


Over-complete basis set
F
F
OH
O
O
N
N
N
N

Processing variables

Processing variables e.g. heat treatment time and temperature have significant impact on a material

Additive manufacturing processing variables can change material properties without requiring recertification of the material

Ontology must be sufficiently flexible to encompass these parameters

Widmanstätten pattern in Kamacite meteorite requires cooling over a million years

Esoteric variables

Significant fraction of the value in a material is setting up its mass manufacture to crucial to capture variables unique to a factory

Ontology must be able to adopt bespoke parameters

How many windows are open in this steel mill?

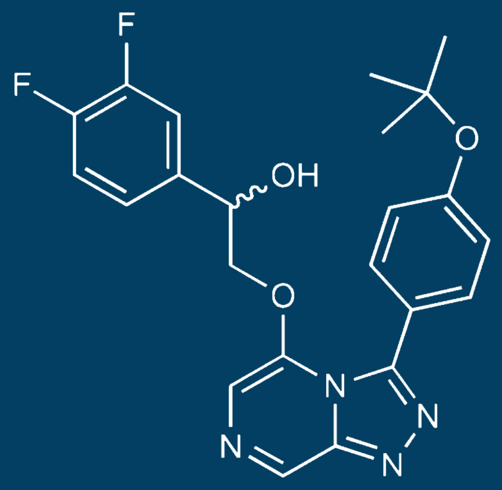
Basis set

Molecular basis set has e.g. 320 components

Over-complete plurality and not orthogonal but contains physically important information for chemists

Framework allows machine learning to exploit property-property relationships

Ingredient has several sub-ingredients, e.g. filler in polymer but some of their sub-ingredients are common e.g. water, which must then be summed over



# O	3	oxygen
# O - H	1	hydroxyl
# C - O - C	2	ether
# C = O	0	carbonyl

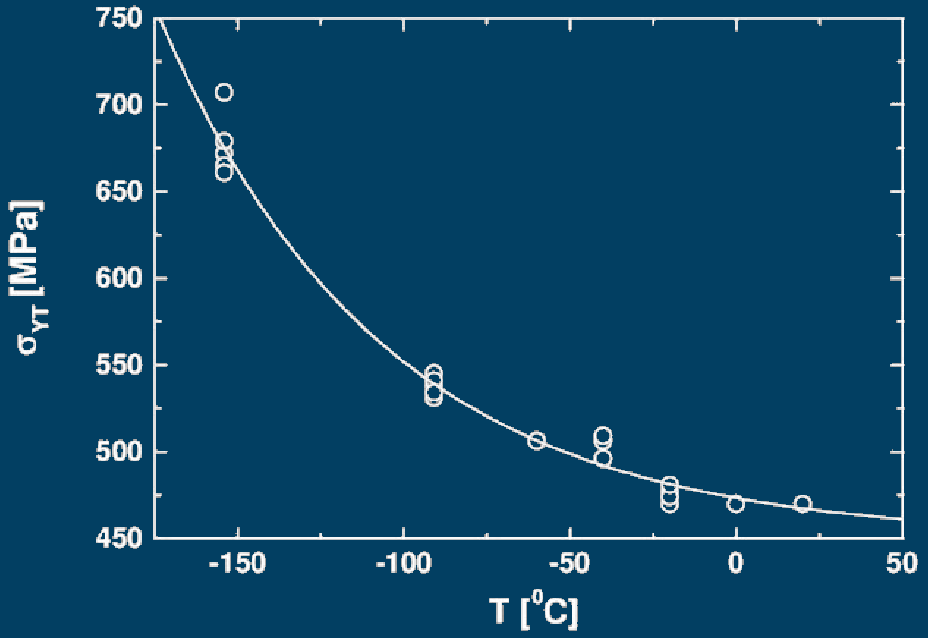
Winner of Open Source Malaria competition
Accepted for publication J. of Medicinal Chem.

Custom data types

Data is often a graph e.g. yield stress vs temperature

Inefficient to store each measurement as a separate material

Helpful to handle such vector data



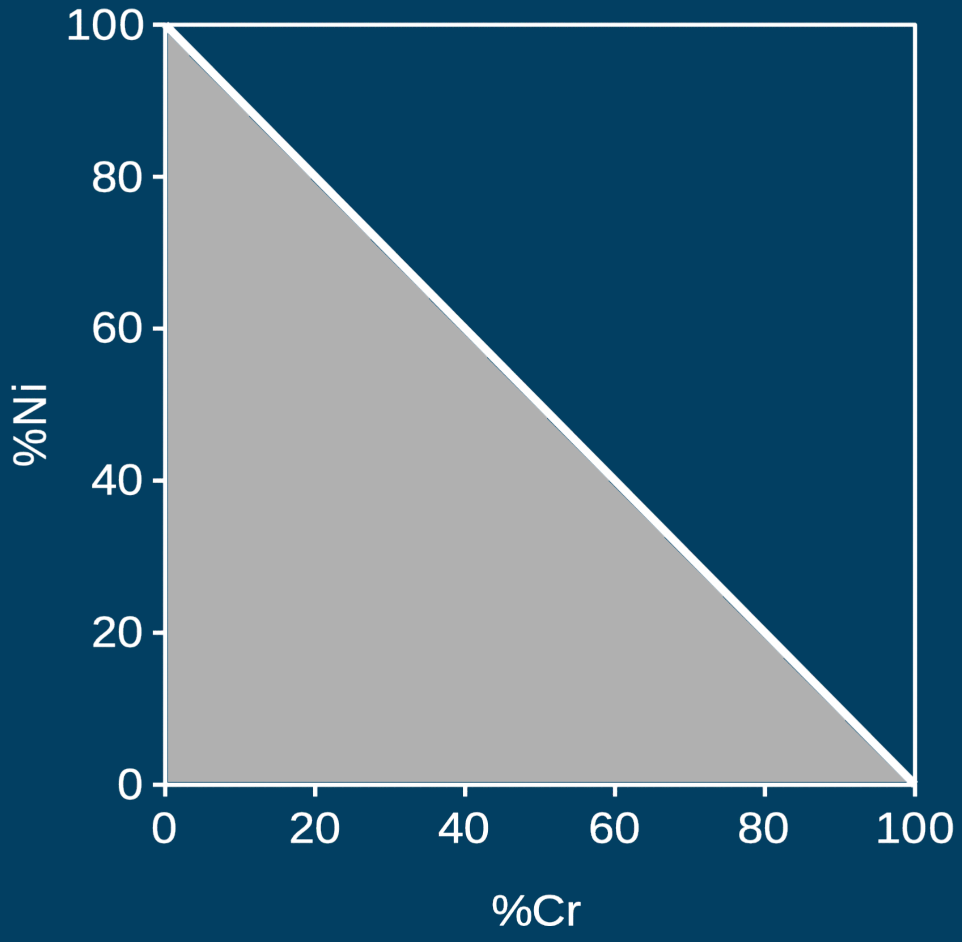
22NiMoCr37
The “EURO” Fracture Toughness Curve.
Technical Research Centre of Finland
K. Wallin, (1998)

Constraints on design space

When designing materials the opportunity space can have non-trivial constraints

Composition usually has a balance element, remaining elements must sum to less than 100%, defining a simplex

In laser deposition volumetric energy density limited but actual parameter is laser power and part volume



Three component Fe-Ni-Cr alloy with Fe as balance element

Security and provenance

Companies confidential about data so ontology should enable access control

Versioning and referencing of data important to track provenance, quality, and give proper credit

Summary

Machine learning is a powerful tool to capture complex relationships

Approach with property-property relationships commercialized by intellegens

Highlight six considerations for development of a materials ontology for industrial applications

- Processing variables
- Esoteric variables
- Over-complete basis set
- Constraints on design space
- Custom data types
- Security and provenance