

自己組織化マップ (SOM) を用いた データ駆動型材料スクリーニング: 分子シミュレーションとデータ科学の狭間

菊川 豪太 / Gota KIKUGAWA*

*¹ 東北大学 流体科学研究所

kikugawa@tohoku.ac.jp

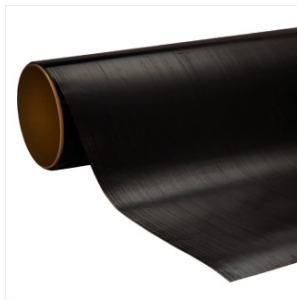


“シミュレーション科学とデータ科学の融合”
July 3, 2020

航空機用構造材料に用いられる架橋高分子材料の開発

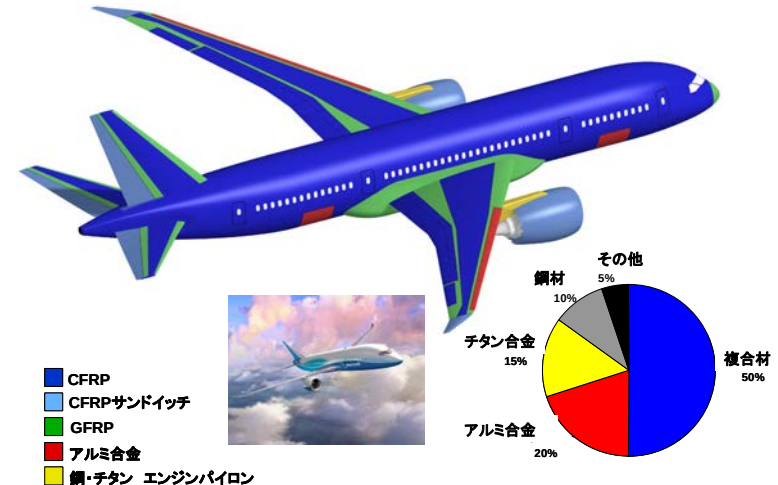
構造材としての架橋高分子材料

- 航空機用構造材料として炭素繊維強化プラスチック(CFRP)複合材の利用が今後益々広がる.
- 架橋高分子は,2種類の材料(エポキシ材料+硬化剤)を混合
⇒架橋反応によってネットワーク形成
- 多数の候補材料から最適な材料を選び(スクリーニング),材料の内部構造や材料特性を効率よく予測および制御することが重要



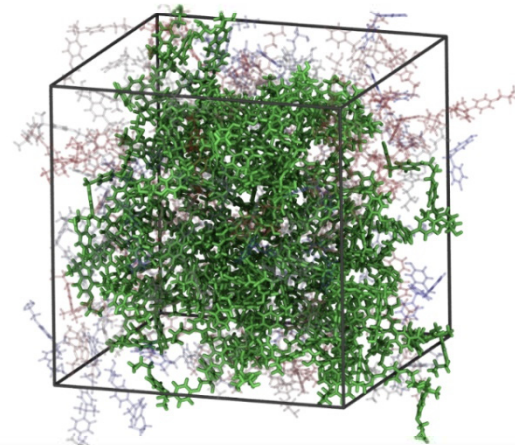
https://www.torayca.com/lineup/product/pro_003.html

プリプレグ(炭素繊維樹脂シート)



<http://www.boeing.com/commercial/787/>

航空機における複合材利用



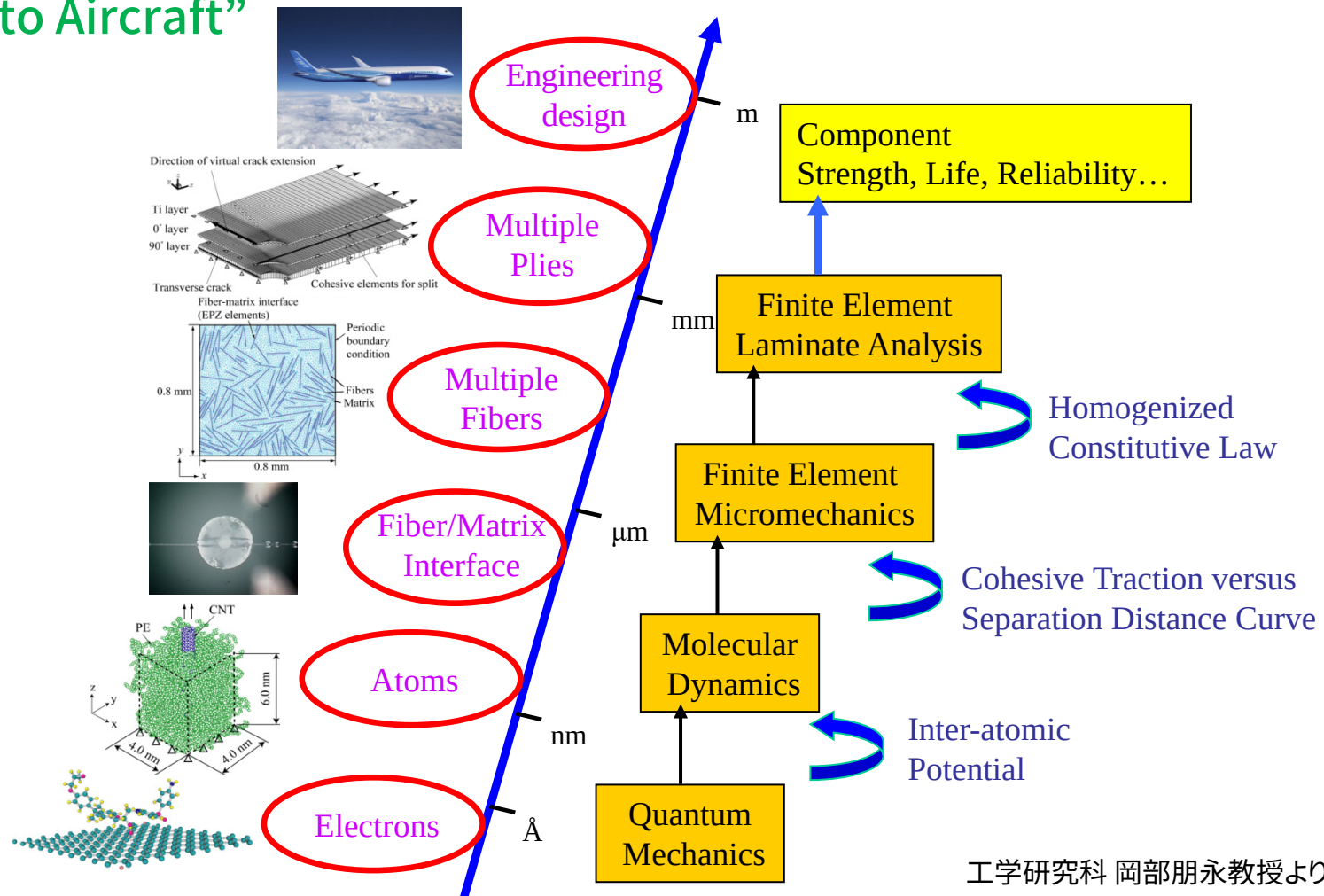
エポキシ樹脂の3次元ネットワーク構造

航空機に使用される複合材料のマルチスケールモデリング

複合材料のマルチスケール数値シミュレーション

サイバー空間での高速な材料物性予測

“Atoms to Aircraft”

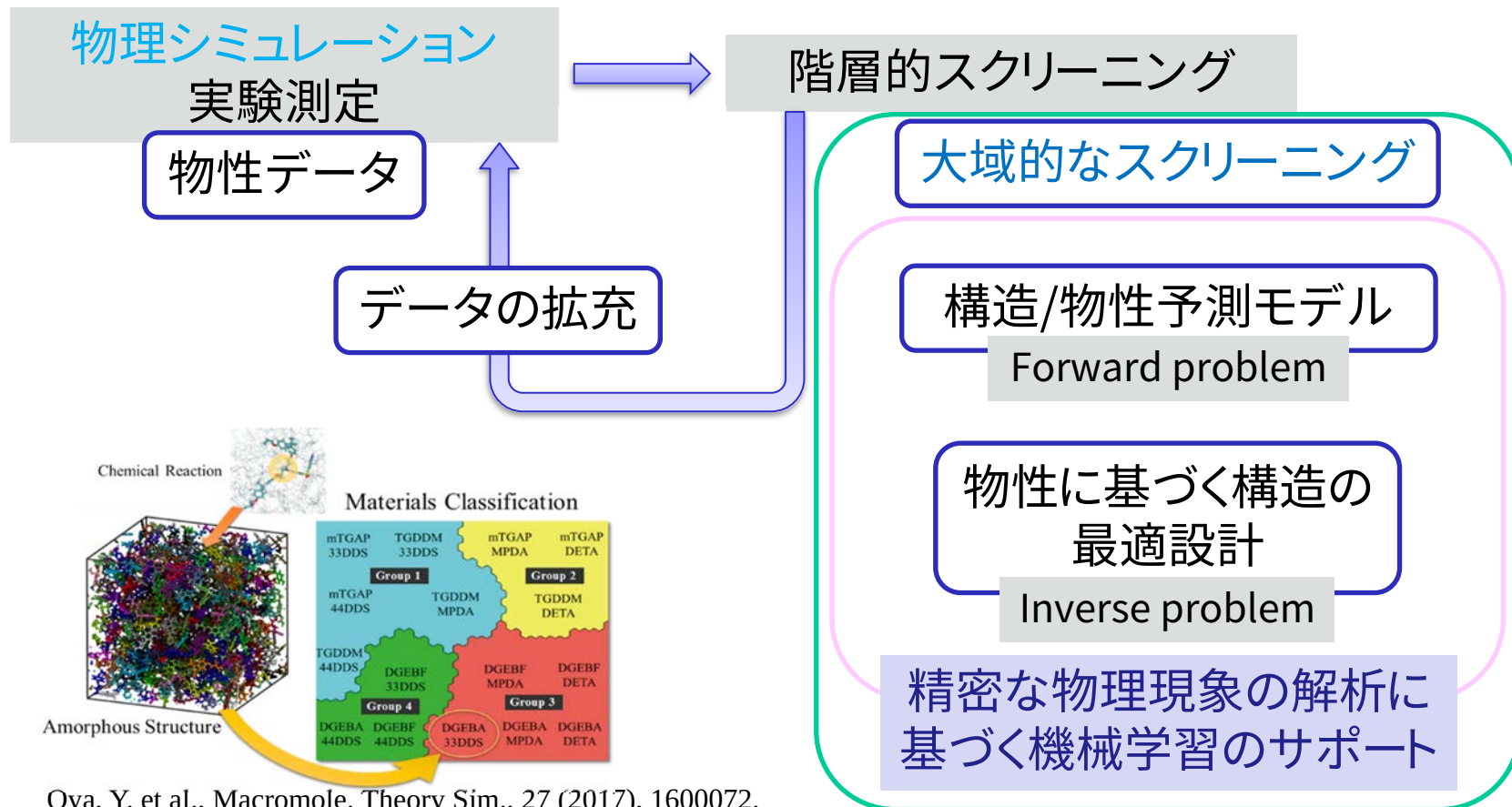


工学研究科 岡部朋永教授より提供

材料インフォマティクスを利用した材料探索

● 材料インフォマティクス (Materials informatics, MI)

- ✓ 大規模な物理シミュレーションとそこから得られる材料物性データに基づくデータマイニングおよび高速材料設計技術の開発



SOM-clustering approach for polymeric materials

ポリマー系MIにおける問題点

高分子系MI研究に存在する(共通の)問題意識

- 本質的にデータ量が少ない問題をどう克服するか？
 - ⇒ データベース
 - ⇒ 分子シミュレーションによる生成
- 高次構造や製造プロセスによる材料のバリエーション(ばらつき)の多さ
 - ⇒ 分岐, 重合度, 分散性
 - ⇒ 結晶, 非晶質(アモルファス)
- 実験データとシミュレーションの差をどう取り扱うか？
 - ⇒ 分子シミュレーションの結果が必ずしも正確というわけではない.

材料探索への機械学習アプローチ

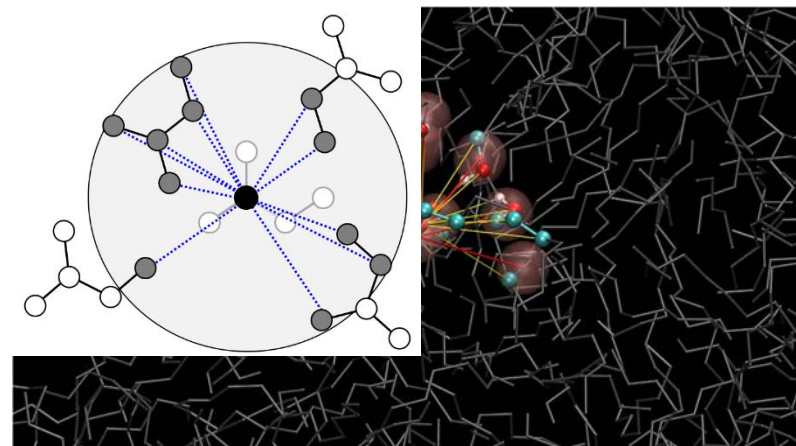
● データ駆動か物理駆動か？

- ✓ 材料探索に**大規模データセット**を利用する,あるいは生成する.
- ✓ 材料物性を支配する要因とは何かを原理原則から調べ(**ボトムアップ的**),それらの組み合わせで所望の物性を有する材料を提案する.

● 物理ベースの機械学習

- ✓ “Scientific” machine learning
- ✓ **Theory-guided** (Physics informed) machine learning
⇒ 解釈問題

Wagner, N. and Rondinelli, M., Frontier Mater., 3 (2016), 28.



Atomic heat path in alcohol liquid

Matsubara, H. et al., Fluid Phase Equilib., Vol. 441, (2017), pp. 24-32.

自己組織化マップ (SOM)

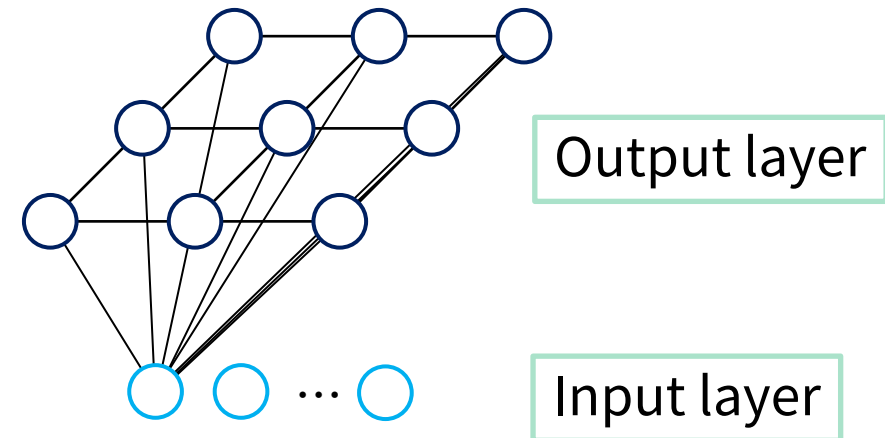
● SOMとは?

* T. Kohonen, "Self-Organizing Maps, third ed.", Springer-Verlag, 2001.

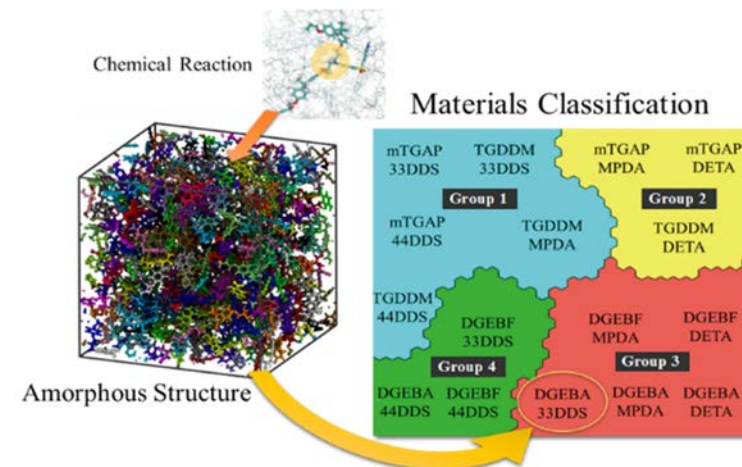
- ✓ **次元圧縮**: 多次元の情報を**近傍関係を維持しながら**, 低次元に非線形写像
- ✓ 2層のニューラルネットにより構成される教師なし学習
- ✓ **視覚的にデータの相互関係を理解しやすい**.

Application

- ビジネスマーケティングにおける消費志向の抽出
- ゲノム解析
- 航空機翼の形状最適化^{*1}
- 架橋高分子樹脂の探索^{*2}



- ✓ 多次元の物性を持つ多数の材料

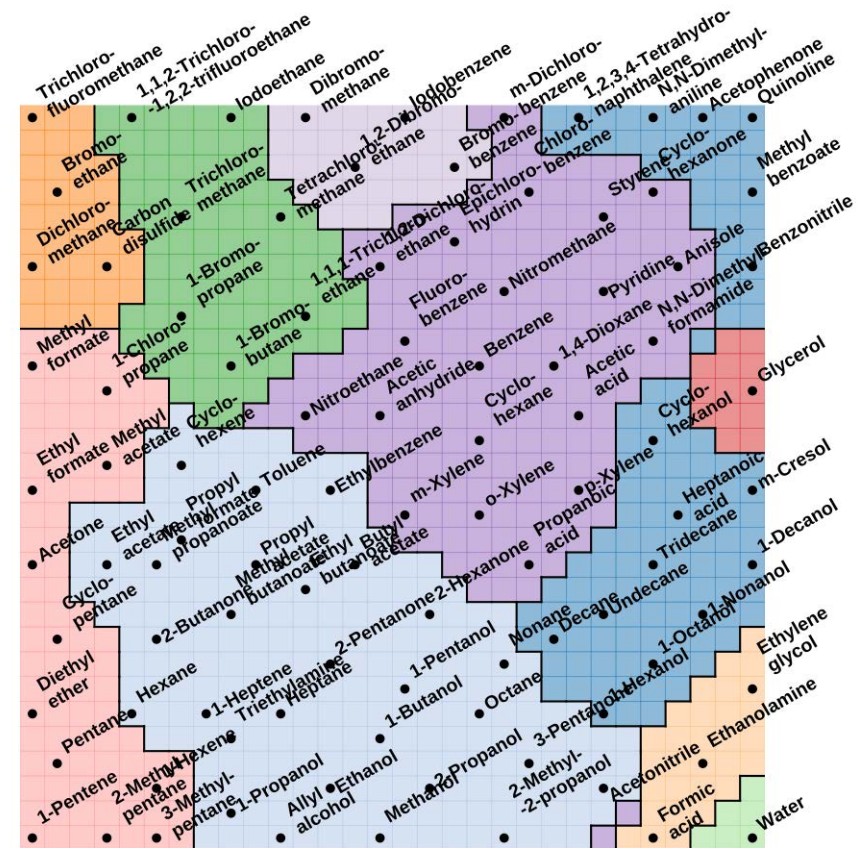


^{*1} Shimoyama, K. and Kamisori, K., J. Aircraft, 54 (2017), pp. 1317-1327.

^{*2} Oya, Y. et al., Macromole. Theory Sim., 27 (2017), 1600072.

SOMークラスタリング解析

- 98種類の液体の熱物性をデータベース等から取得
- 入力データとして8種類の物性: 密度, 比熱, 融点, 沸点, 蒸気圧, 表面張力, 粘性, 熱伝導率
- 化学構造のデータは含まない

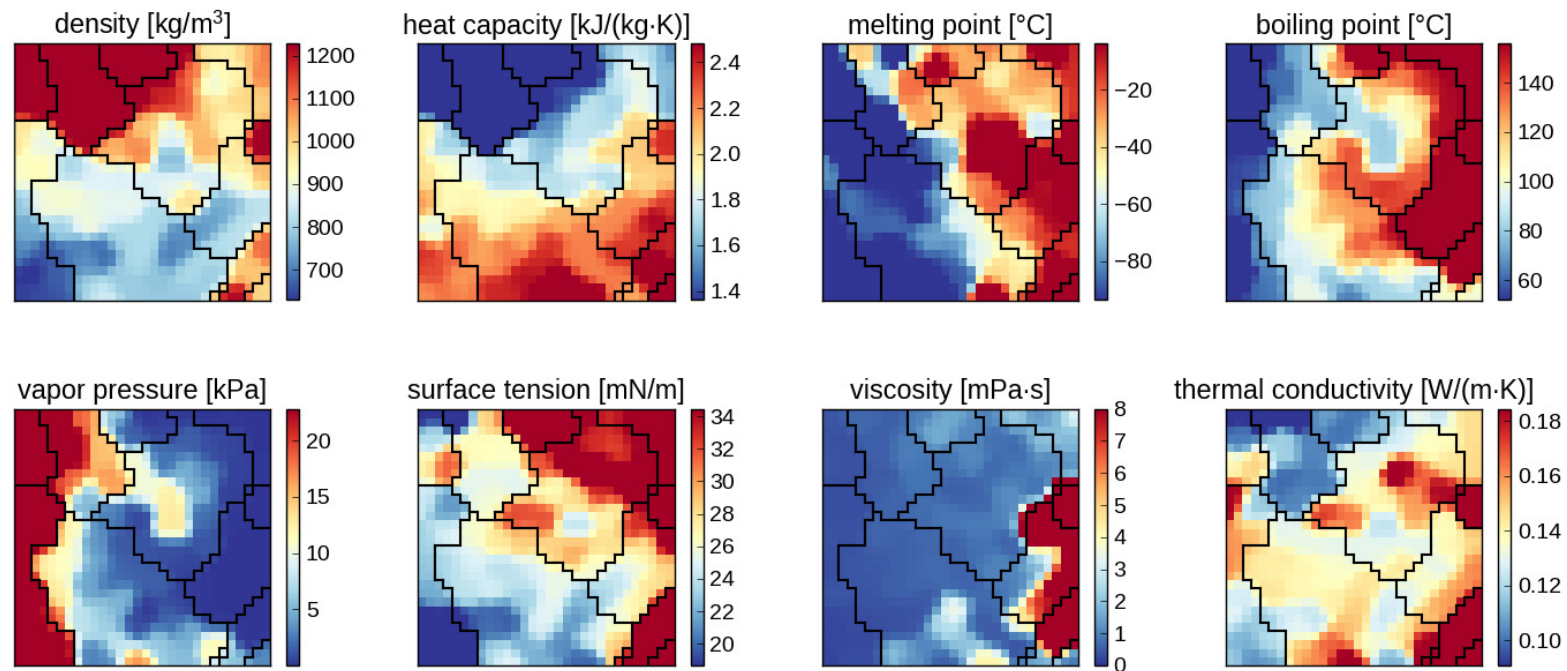


SOM positioning map

- 多数の液体が官能基によって特徴付けられる分子種に分類されている。

* Kikugawa, G. et al., Chem. Phys. Lett., 728 (2019), pp. 109-114.

SOMマップ上における熱物性表示



Heat map of thermophysical properties

- 一目で各材料・各物性の相関関係・相反関係がわかる.
- 例) 冷媒に適した材料種: **高い熱伝導率**かつ**低い粘性**, **低密度**
⇒メタノール, ギ酸, ギ酸メチルなど
(ただしここでは毒性や可燃性などは評価していない)

高精度・高速化へ向けた今後の展開

クラスタ分析に対する量子アニーリングの適用

- 量子アニーリングマシンの原理であるイジングモデルの基底状態計算をクラスタ分析として定式化



<https://www.dwavesys.com/>

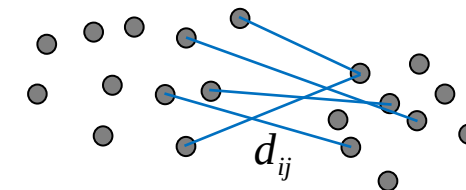
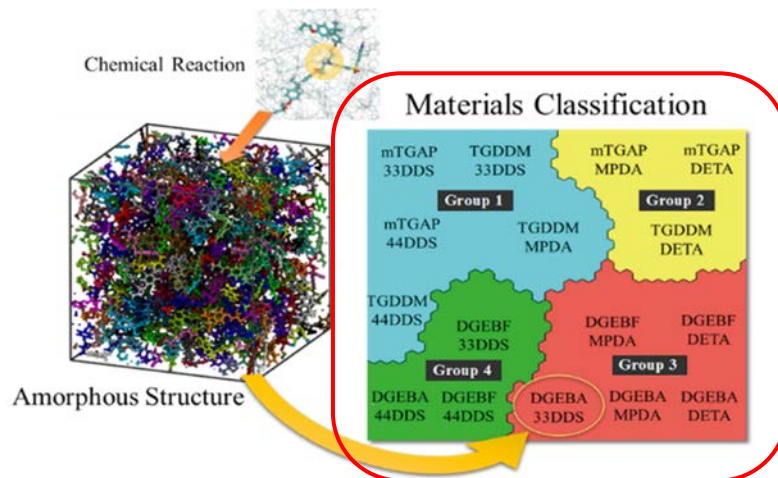
$$H = -\sum_{i<j} J_{ij} \sigma_i \sigma_j$$



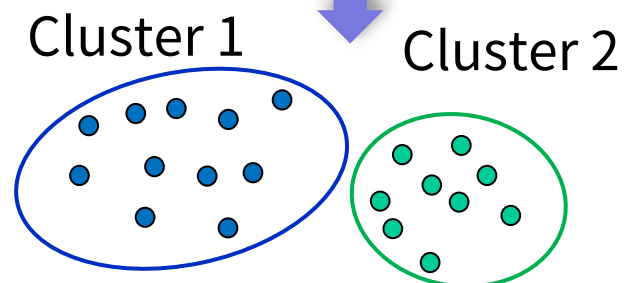
$$D = \frac{1}{2} \sum_{i \neq j} d_{ij} (1 - \sigma_i \sigma_j)$$

イジングモデルのハミルトニアン

クラスタ分析の多次元クラスタ間距離



異なるクラスタに属する要素間の距離の和を最大化



- SOM学習結果のグルーピングに活用