Self-Organizing Map Technique: Data Visualization through Dimensionality Reduction

Yutaka Oya (Ph.D.)
Assistant Professor
Department of Aerospace Engineering, Tohoku University
*Currently Department of Chemistry, University of Washington

Traditional Way of Presenting Data

Properties

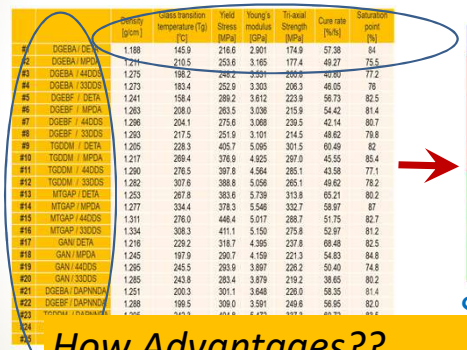
	Density [g/cm³]	Glass transition temperature (Tg) [℃]	Yield Stress [MPa]	Young's modulus [GPa]	Tri-axial Strength [MPa]	Cure rate [%/fs]	Saturation point [%]
#1	DGEBA / DETA	1.188	145.9	216.6	2.901	174.9	57.38	84
#2	DGEBA / MPDA	1.211	210.5	253.6	3.165	177.4	49.27	75.5
#3	DGEBA / 44DDS	1.275	198.2	248.2	3.531	200.8	40.80	77.2
#4	DGEBA / 33DDS	1.273	183.4	252.9	3.303	206.3	46.05	76
#5	DGEBF / DETA	1.241	158.4	289.2	3.612	223.9	56.73	82.5
#6	DGEBF / MPDA	1.263	208.0	263.5	3.036	215.9	54.42	81.4
#7	DGEBF / 44DDS	1.296	204.1	275.6	3.068	239.5	42.14	80.7
#8	DGEBF / 33DDS	1.293	217.5	251.9	3.101	214.5	48.62	79.8
#9	TGDDM / DETA	1.205	228.3	405.7	5.095	301.5	60.49	82
#10	TGDDM / MPDA	1.217	269.4	376.9	4.925	297.0	45.55	85.4
#11	TGDDM / 44DDS	1.290	276.5	397.8	4.564	285.1	43.58	77.1
#12	TGDDM / 33DDS	1.282	307.6	388.8	5.056	265.1	49.62	78.2
#13	MTGAP / DETA	1.253	267.8	383.6	5.739	313.8	65.21	80.2
#14	MTGAP / MPDA	1.277	334.4	378.3	5.546	332.7	58.97	87
#15	MTGAP / 44DDS	1.311	276.0	446.4	5.017	288.7	51.75	82.7
#16	MTGAP / 33DDS	1.334	308.3	411.1	5.150	275.8	52.97	81.2
#17	GAN/ DETA	1.216	229.2	318.7	4.395	237.8	68.48	82.5
#18	GAN / MPDA	1.245	197.9	290.7	4.159	221.3	54.83	84.8
#19	GAN / 44DDS	1.295	245.5	293.9	3.897	226.2	50.40	74.8
#20	GAN / 33DDS	1.285	243.8	283.4	3.879	219.2	38.65	80.2
#21	DGEBA / DAPNNDA	1.251	200.3	301.1	3.648	228.0	58.35	81.4
#22	DGEBF / DAPNNDA	1.288	199.5	309.0	3.591	249.6	56.95	82.0
#23	TGDDM / DAPNNDA	1.295	242.3	404.8	5.472	337.3	60.72	83.5
#24	MTGAP / DAPNNDA	1.332	241.7	423.8	5.332	314.0	60.12	81.0
#25	GAN / DAPNNDA	1.307	209.3	389.5	4.719	291.2	51.30	84.4

Materials

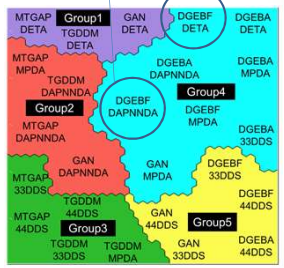
Date-Base

Self Organizing Maps

Properties



Materials



Self-Organization!

How Advantages??
Let's Compare to the Other Teqniques

Visualization Techniques

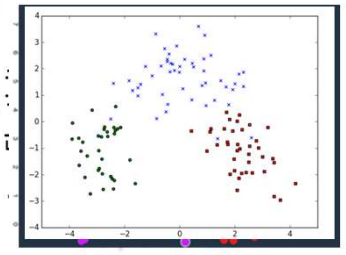
	Density [g/cm³]	Glass transition temperature (Tg) [℃]	Yield Stress [MPa]	Young's modulus [GPa]	Tri-axial Strength [MPa]	Cure rate [%/fs]
#1	1.188	145.9	216.6	2.901	174.9	57.38
#2	1.211	210.5	253.6	3.165	177.4	49.27
#3	1.275	198.2	248.2	3.531	200.8	40.80
#4	1.273	183.4	252.9	3.303	206.3	46.05
#5	1.241	158.4	289.2	3.612	223.9	56.73
#6	1.263	208.0	263.5	3.036	215.9	54.42
#7	1.296	204.1	275.6	3.068	239.5	42.14
#8	1.293	217.5	251.9	3.101	214.5	48.62
#9	1.205	228.3	405.7	5.095	301.5	60.49
#10	1.217	269.4	376.9	4.925	297.0	45.55
#11	1.290	276.5	397.8	4.564	285.1	43.58
#12	1.282	307.6	388.8	5.056	265.1	49.62

Data Set

Precise (quantitative) Information

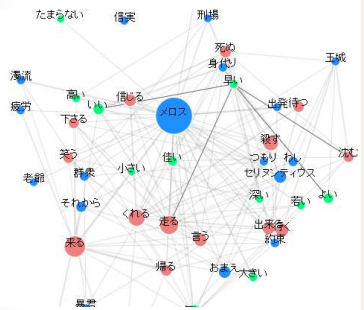
Data Set

Understanding Features

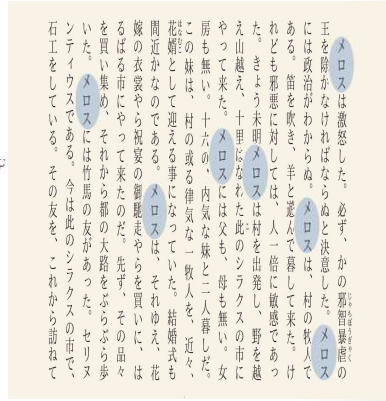


Co-Occurrence Network

Connectivity and Frequency of Each Word



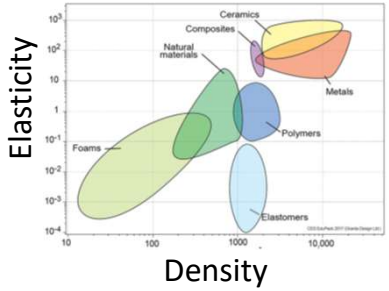
https://textmining.userlocal.jp/home/result/0c4e44ad98ed769ef4a77176009ba155



メロスは激怒した。必ず、かの邪智暴虐の王を除かなければならぬと決意した。メロスには政治がわからぬ。メロスは、村の牧人である。笛を吹き、羊と遊んで暮して来た。けれども邪悪に対しては、人一倍に敏感であった。きょう未明メロスは村を出発し、野を越え山越え、十里になれた此のシラクスの市にやって来た。メロスには父も、母も無い。女房も無い。十六の、内気な妹と二人暮しだ。この妹は、村の或る律気な一牧人を、近々、花婿として迎える事になっていた。結婚式も間近かなのである。メロスは、それゆえ、花嫁の衣裳やら祝宴の御馳走やらを買いに、はるばる市にやって来たのだ。先ず、その品々を買い集め、それから都の大路をぶらぶら歩いた。メロスには竹馬の友があった。セリヌンティウスである。今は此のシラクスの市で、石工をしている。その友を、これから訪ねて

Ashby Maps

Material Distributes on Two Dimensional Space

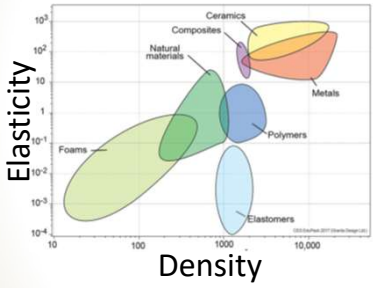


http://teachingresources.grantadesign.com/mikecharts/young-modulus-vs-density

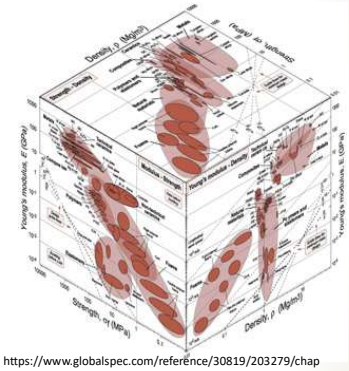
Understand the relationship between Materials and Properties

Ashby Maps

Material Distributes on Two Dimensional Space



http://teachingresources.grantadesign.com/mikecharts/young-modulus-vs-density

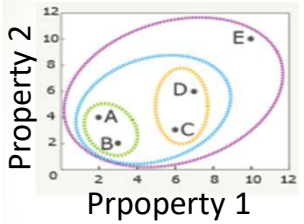


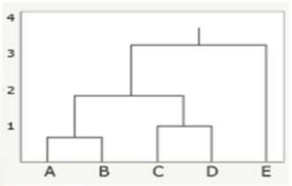
https://www.globalspec.com/reference/30819/203279/chapter-4-material-property-charts

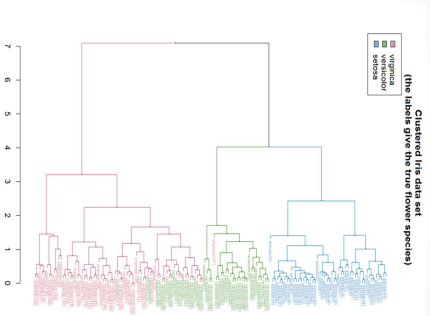
Understand the relationship between Materials and Properties

Hierarchical Clustering

1-D Sorting depends on Similarities



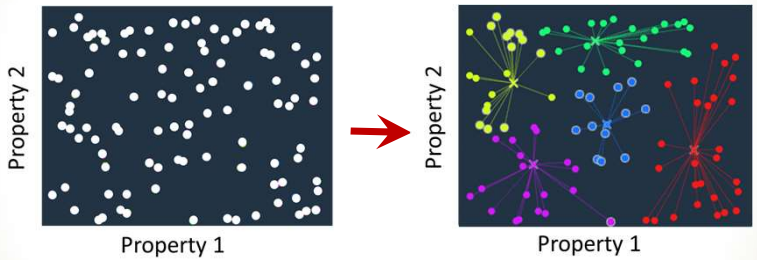




https://about.yahoo.co.jp/info/bigdata/special/2016/05/app/

K-Means Method

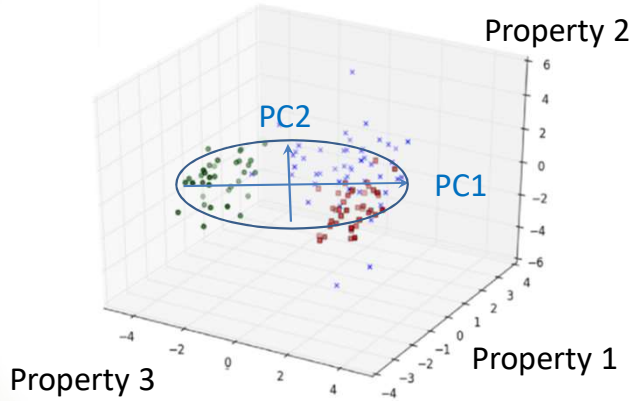
Clustering method for Dispersed data



http://tech.nitoyon.com/ja/blog/2009/04/09/kmeans-visualise/

Principal component analysis

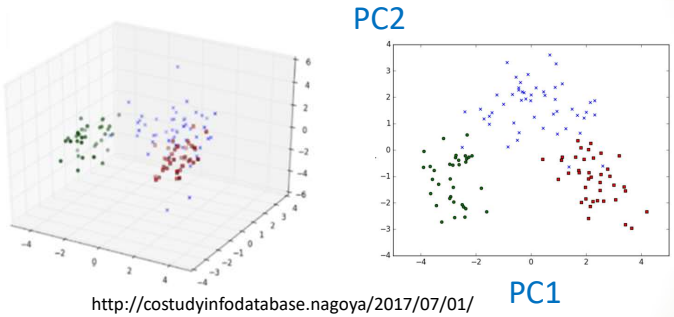
Dimensionally Reduction



http://costudyinfodatabase.nagoya/2017/07/01/

Principal component analysis

Dimensionally Reduction



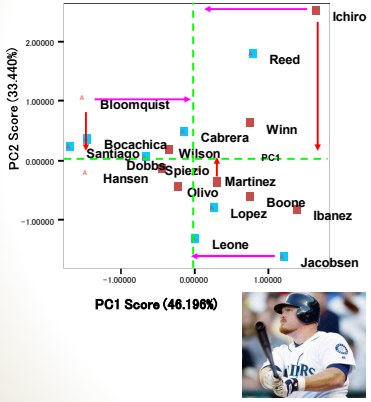
http://costudyinfodatabase.nagoya/2017/07/01/

3Dimension → 2Dmension

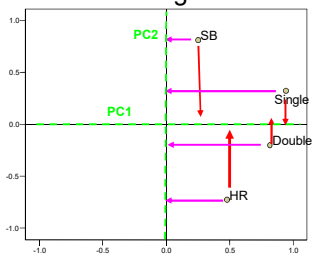
Principal component analysis

Dimensionally Reduction

Score plot



Loading Plot



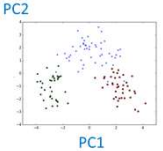
Ichiro : Good baseball player with speed

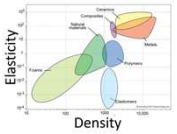
Jacobsen : Powerful but slow

Martinez : Good timing to retire

Summery of Visualization Techniques

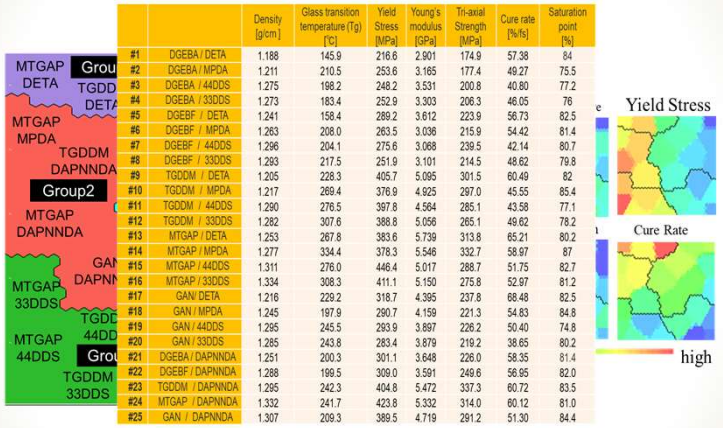
- Dimensionally Reduction
- Clustering
- 1 to 1 Relation between Materials and Properties
- Sorting Depends on Similarity





Self-Organizing Maps

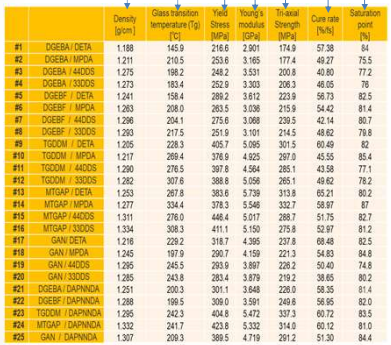
Overview of Self Organizing Maps (SOM)

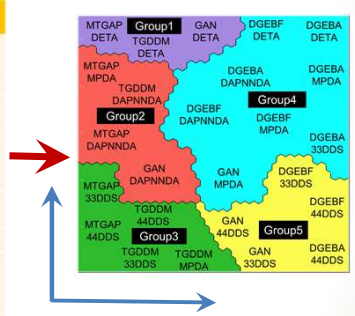


		Density [g/cm³]	Glass transition temperature (Tg) [°C]	Yield Stress [MPa]	Young's modulus [GPa]	Tri-axial Strength [MPa]	Cure rate [%/fs]	Saturation point [%]
#1	DGEBA / DETA	1.188	145.9	216.6	2.901	174.9	57.38	84
#2	DGEBA / MPDA	1.211	210.5	253.6	3.165	177.4	49.27	75.5
#3	DGEBA / 44DDS	1.275	198.2	248.2	3.531	200.8	40.80	77.2
#4	DGEBA / 33DDS	1.273	183.4	252.9	3.303	206.3	46.05	76
#5	DGEBF / DETA	1.241	158.4	289.2	3.612	223.9	56.73	82.5
#6	DGEBF / MPDA	1.263	208.0	263.5	3.036	215.9	54.42	81.4
#7	DGEBF / 44DDS	1.296	204.1	275.6	3.068	239.5	42.14	80.7
#8	DGEBF / 33DDS	1.293	217.5	251.9	3.101	214.5	48.62	79.8
#9	TGDDM / DETA	1.205	228.3	405.7	5.095	301.5	60.49	82
#10	TGDDM / MPDA	1.217	269.4	376.9	4.925	297.0	45.55	85.4
#11	TGDDM / 44DDS	1.290	276.5	397.8	4.564	285.1	43.58	77.1
#12	TGDDM / 33DDS	1.282	307.6	388.8	5.056	265.1	49.62	78.2
#13	MTGAP / DETA	1.253	267.8	383.6	5.739	313.8	65.21	80.2
#14	MTGAP / MPDA	1.277	334.4	378.3	5.546	332.7	58.97	87
#15	MTGAP / 44DDS	1.311	276.0	446.4	5.017	288.7	51.75	82.7
#16	MTGAP / 33DDS	1.334	308.3	411.1	5.150	275.8	52.97	81.2
#17	GAN/ DETA	1.216	229.2	318.7	4.395	237.8	68.48	82.5
#18	GAN / MPDA	1.245	197.9	290.7	4.159	221.3	54.83	84.8
#19	GAN / 44DDS	1.295	245.5	293.9	3.897	226.2	50.40	74.8
#20	GAN / 33DDS	1.285	243.8	283.4	3.879	219.2	38.65	80.2
#21	DGEBA / DAPNNDA	1.251	200.3	301.1	3.648	226.0	58.35	81.4
#22	DGEBF / DAPNNDA	1.288	199.5	309.0	3.591	249.6	56.95	82.0
#23	TGDDM / DAPNNDA	1.295	242.3	404.8	5.472	337.3	60.72	83.5
#24	MTGAP / DAPNNDA	1.332	241.7	423.8	5.332	314.0	60.12	81.0
#25	GAN / DAPNNDA	1.307	209.3	389.5	4.719	291.2	51.30	84.4

Overview of SOM

7 Properties (7dimension)

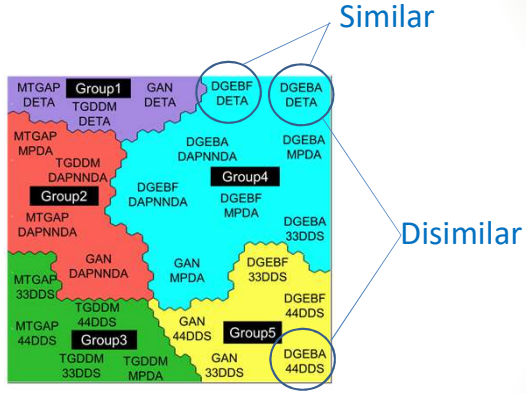




- Dimensionally Reduction: 7D→2D

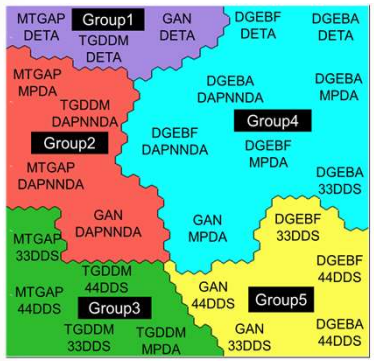
Overview of SOM

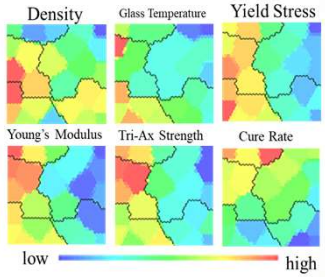
Color Clustering



- Clustering and Sorting depending on Similarity

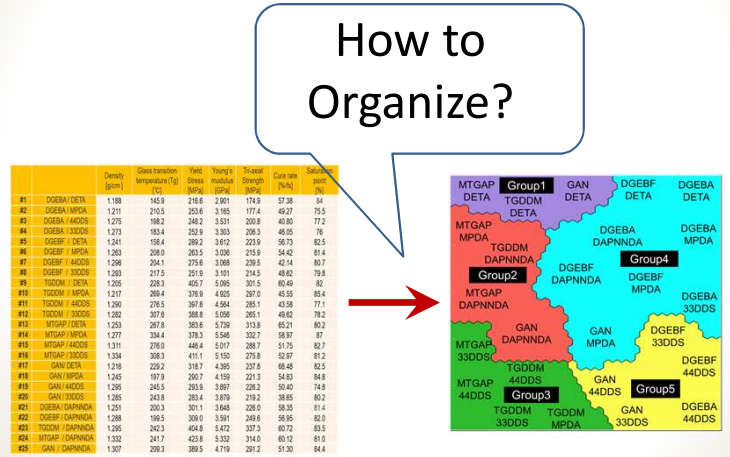
Overview of SOM





- 1 to 1 relation between Materials and Properties

Algorithm of Self-Organization



Algorithm of Self-Organization : Data Vectorization

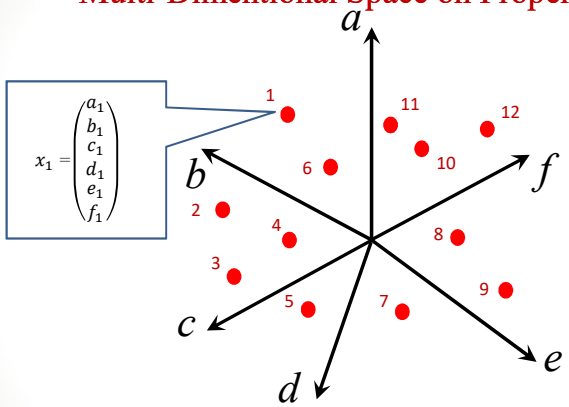
	Density [g/cm³]	Glass transition temperature (Tg) [°C]	Yield Stress [MPa]	Young's modulus [GPa]	Tri-axial Strength [MPa]	Cure rate [%/fs]
#1	1.188	145.9	216.6	2.901	174.9	57.38
#2	1.211	210.5	253.6	3.165	177.4	49.27
#3	1.275	198.2	248.2	3.531	200.8	40.80
#4	1.273	183.4	252.9	3.303	206.3	46.05
#5	1.241	158.4	289.2	3.612	223.9	56.73
#6	1.263	208.0	263.5	3.036	215.9	54.42
#7	1.296	204.1	275.6	3.068	239.5	42.14
#8	1.293	217.5	251.9	3.101	214.5	48.62
#9	1.205	228.3	405.7	5.095	301.5	60.49
#10	1.217	269.4	376.9	4.925	297.0	45.55
#11	1.290	276.5	397.8	4.564	285.1	43.58
#12	1.282	307.6	388.8	5.056	265.1	49.62

$$x_i = \begin{pmatrix} a_i \\ b_i \\ c_i \\ d_i \\ e_i \\ f_i \end{pmatrix}$$

density
Glass transition temperature
Yield Stress
Young's modulus
Tri-axial Strength
Cure rate

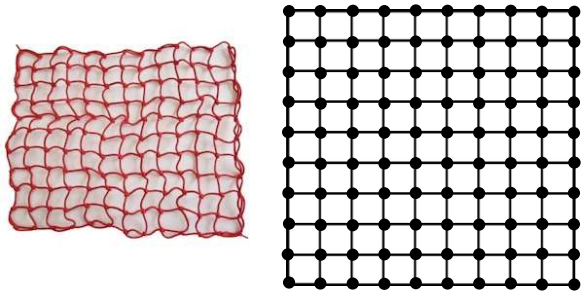
i (1 ~ 12) :Materials $a \sim f$:Properties

Algorithm of SOM : Multi-Dimentional Space on Properties



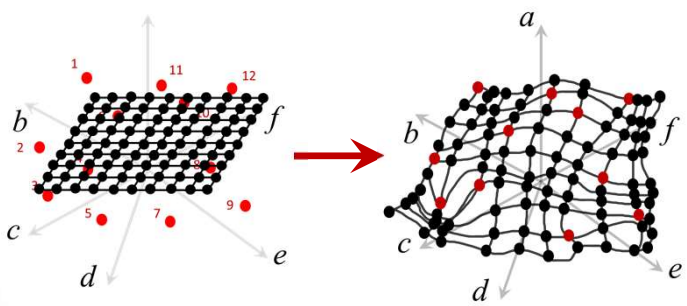
$a \sim f$:Properties

Two Dimensional Mesh Structure

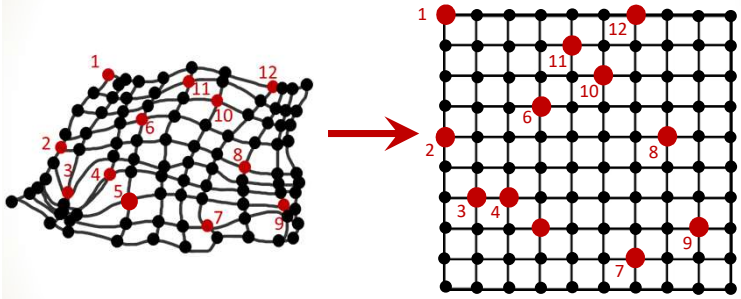


Each connection can deform

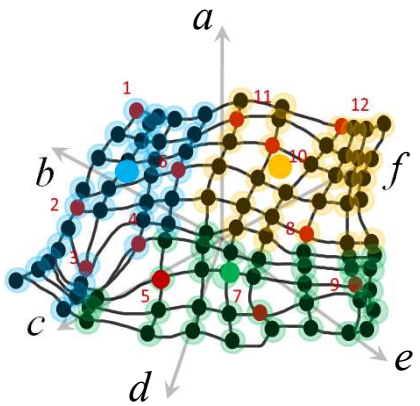
Deformation of the Meshes



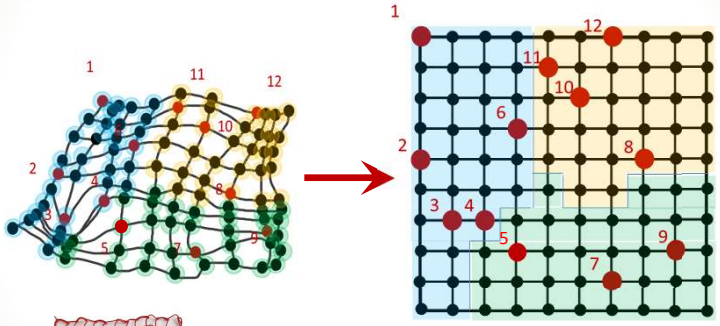
Dimensionality Reduction

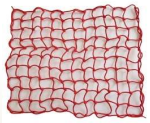


Clusterization of Materials



Catch and Release



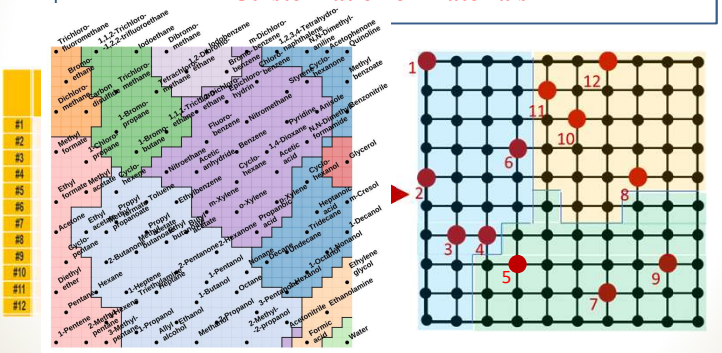


Release Elastic Energy !!

Dimensionality reduction

(High dimension data is projected onto low dimension.)

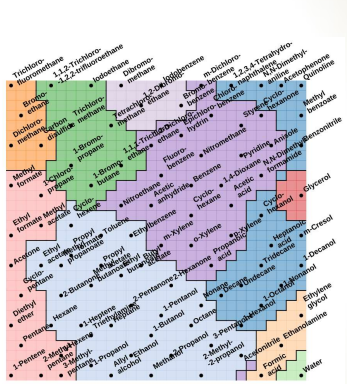
Culsterization of Materials



Application of SOM: Liquid Materials

molecular weight	melting point	boiling point	density	viscosity	heat capacity	vapor pressure	surface tension	thermal conductivity

Acetic acid	C2H4O2
Acetic anhydride	C4H6O3
Acetone	C3H6O
Acetonitrile	C2H3N
Acetophenone	C8H8O
Allyl alcohol	C3H6O
Anisole	C7H8O
Benzene	C6H6

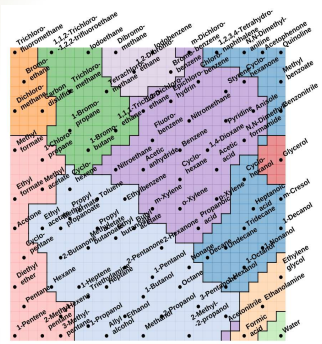


8 Properties

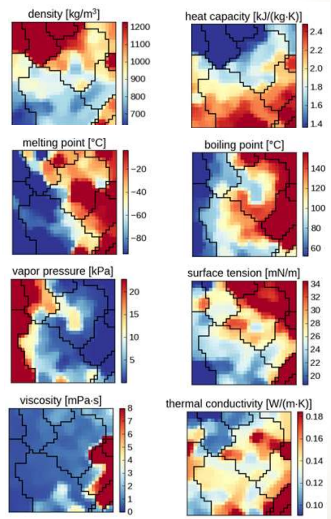
80 Materials

G.Kikugawa, et al (2018)

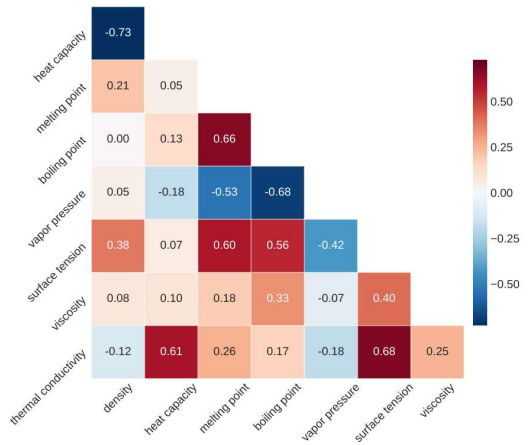
Application of SOM: Liquid Materials



Heat map of thermophysical properties

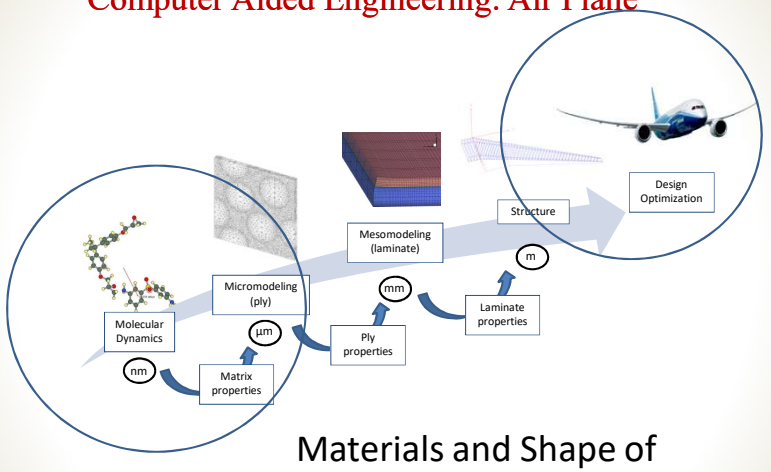


Application of SOM: Liquid Materials



- Overview of correlative relationship between two properties

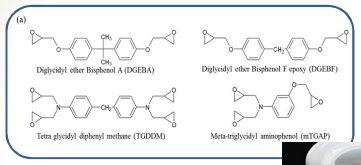
Computer Aided Engineering: Air Plane



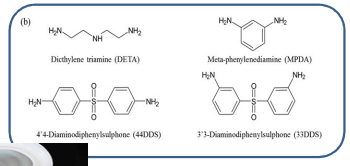
Materials and Shape of Air Plane Wing

Computer Aided Engineering: Air Plane

Base Resin

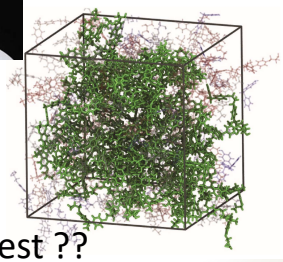


Curing Agent



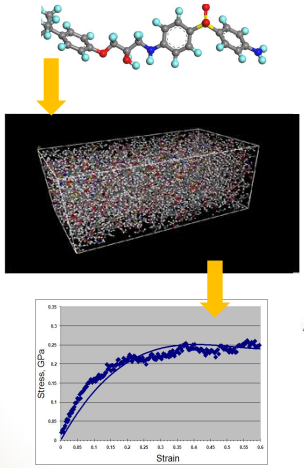
Mixing
Chemical Reaction

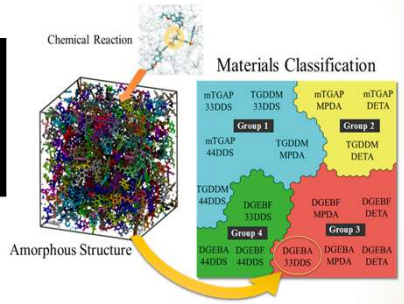
Cross-Linked!!



Which Combination is the best ??

Optimization of Thermosetting Resins

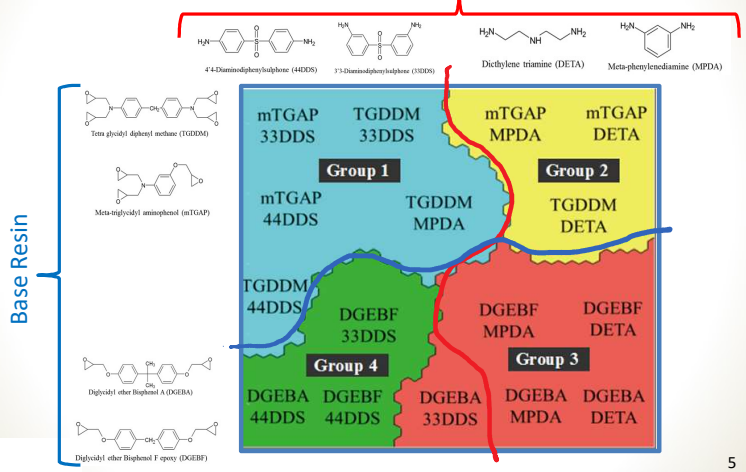




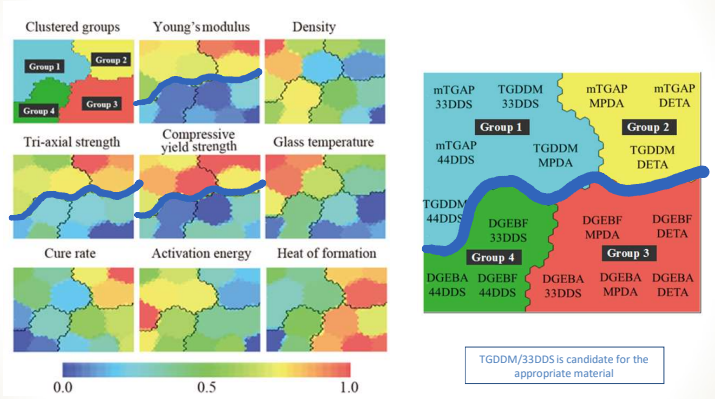
Y.Oya et al (2017)

Optimization of Thermosetting Resins

Curing Agents

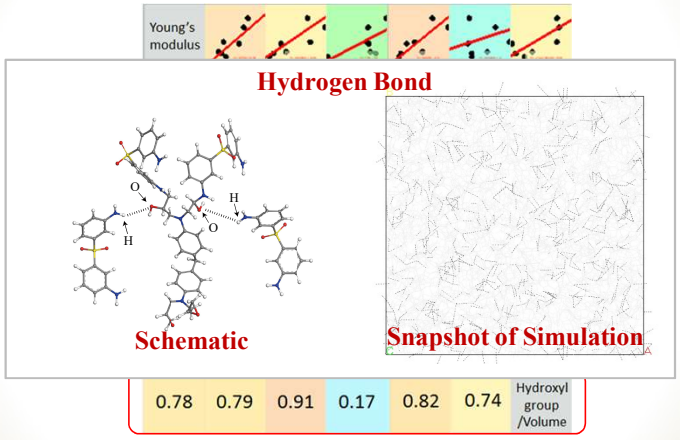


Optimization of Thermosetting Resins



Materials in Group1, and Group 2 have higher Mechanical Properties

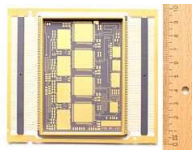
Optimization of Thermosetting Resins



Density of Hydrogen Bond Affect Stiffness and Strength

Working on Project: Optimization of Ceramic





Thermoelectric Figure of Merit

$$FM = \frac{\sigma S^2}{\alpha}$$

Thermal shock

$$Rs = \frac{\lambda \sigma_B}{\alpha E}$$

σ :electric conductivity

S :Seebeck coefficient

α :thermal conductivity

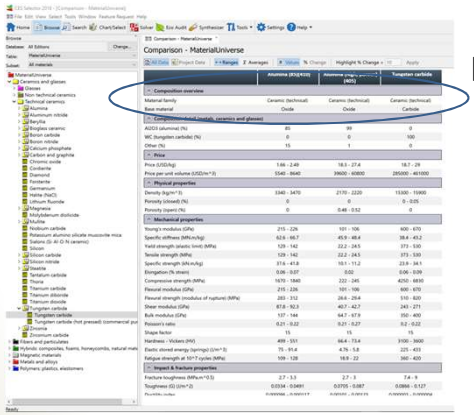
σ_B :mean bending strength

λ :thermal expansion coefficient

E :elastic modulus

Working on Project: Optimization of Ceramic

GRANTA
MATERIAL INTELLIGENCE



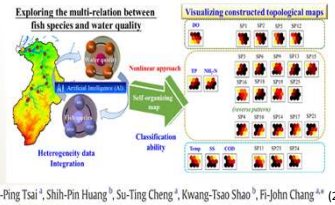
Huge Data Base

Over6000
Ceramics

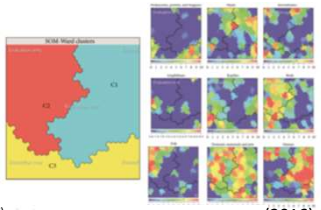
Conclusions

SOM has many advantages to grasp important features of the data through:

- Dimensionally Reduction
- Clustering
- 1 to 1 Relation between Materials and Properties
- Sorting Depends on Similarity



Wen-Ping Tsai [a], Shih-Pin Huang [b], Su-Ting Cheng [a], Kwang-Tsao Shao [b], Fi-John Chang [a,*] (2017)



Nurettin Yorek, Ilker Ugulu, and Halil Aydin (2016)