

---

# Standardization of Materials R&D Data Schema and Vocabulary

Kwang-Ryeol Lee ([krlee@kist.re.kr](mailto:krlee@kist.re.kr))

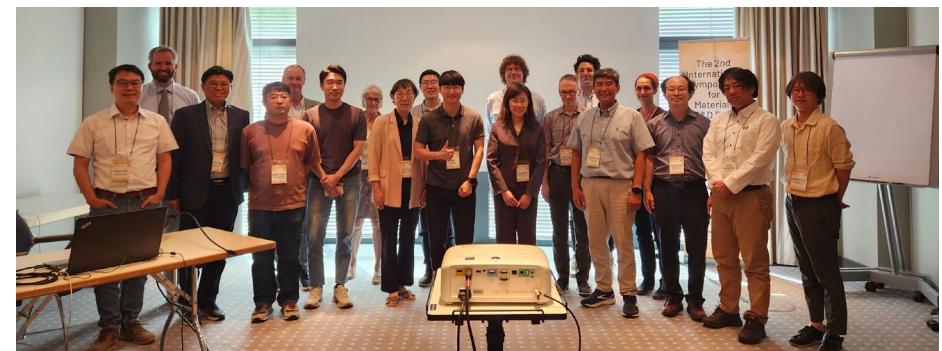
Computational Science Research Center  
Korea Institute of Science and Technology

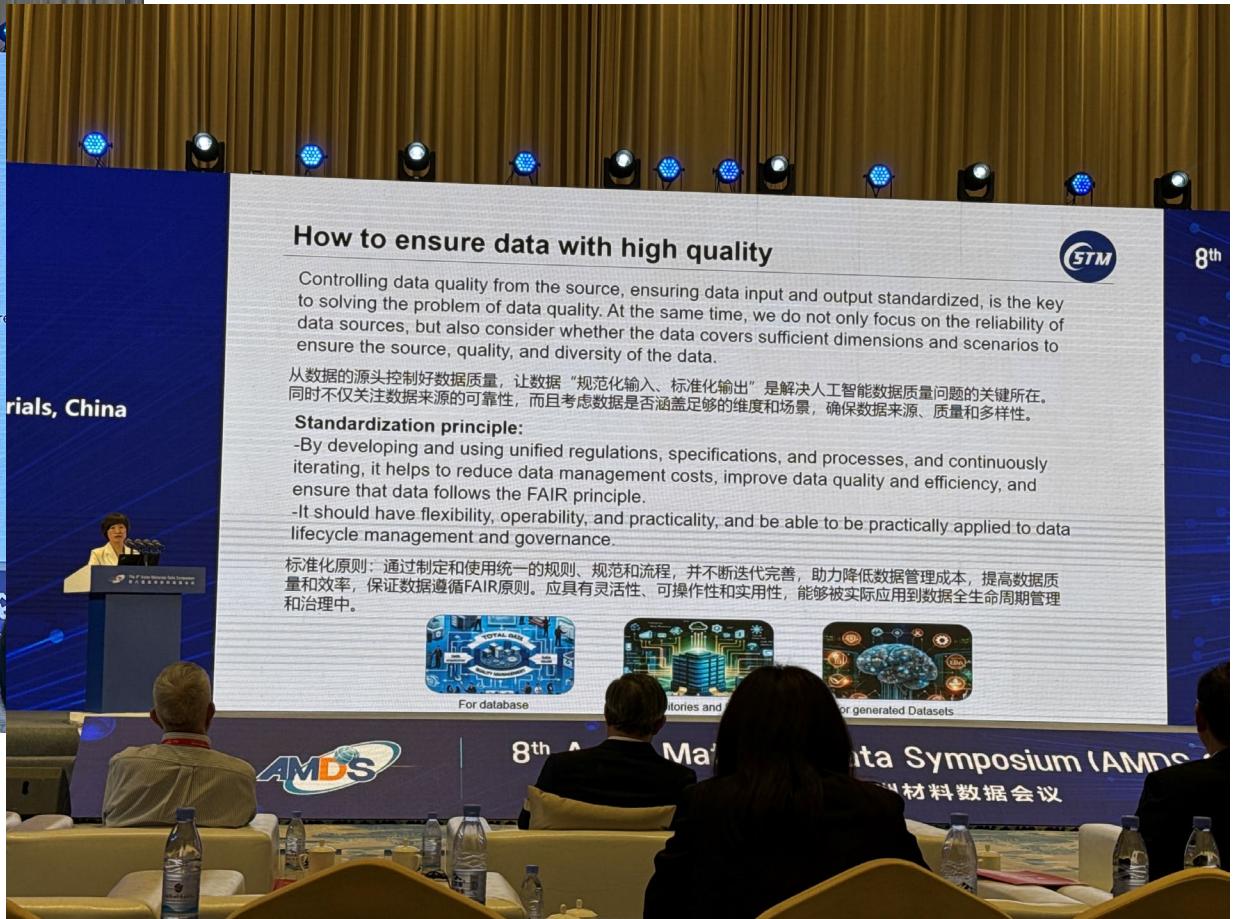
Materials R&D Data Standard Expert Committee  
National Center for Materials Research Data

# Acknowledgement

- National Center for Materials Research Data
- Contributors

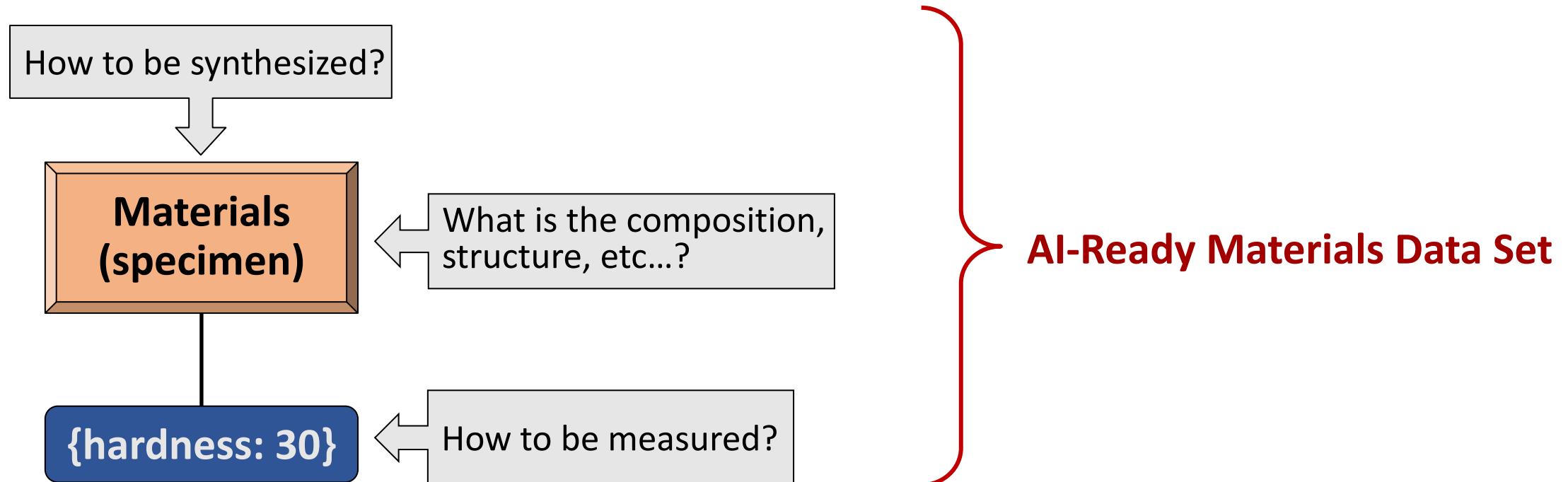
허용학 (KRISS), 장현주 (KRICT), 한승우 (서울대), 김용주 (국민대), 심형석 (KRISS), 한상수 (KIST), 신호선 (KRISS), 오창석 (KIMS), 이정훈 (KIST), 이병주 (KIST), 김지한 (KAIST), 김현유 (충남대), 서동화 (KAIST), 임경근 (KRISS), 이동화 (POSTECH), 박선희 (KRISS), 김수현 (KIMS), 장효선 (KIMS), 김정한 (한밭대), 김정환 (한밭대), 안재평 (KIST), 김낙균 (KIST), 박수형 (KIST) 제승근 (전남대), 명승택 (세종대), 권용우 (홍익대), 김범수 (KRICT), 류권상 (아이피스), Eva Unger (Humboldt U), 이예리 (KRICT), 정증현 (KIST), 최환진 (KBSI), 권기창 (KRISS), 방기훈 (KIST), 이정현 (고려대), 유재상 (KIST), 김재석 (KRISS), 이우 (KRISS), 이상한 (GIST), 정인 (서울대), 김일호 (철도대), 손재성 (POSTECH), 이규형 (연세대), 유태수 (충북대), 이상한 (GIST), 송재용 (POSTECH), 김선태 (전북대), 이관영 (고려대), 이병주 (POSTECH), 임영복 (KIMS), 차필령 (국민대), 최우진 (KRICT), 한홍남 (서울대)



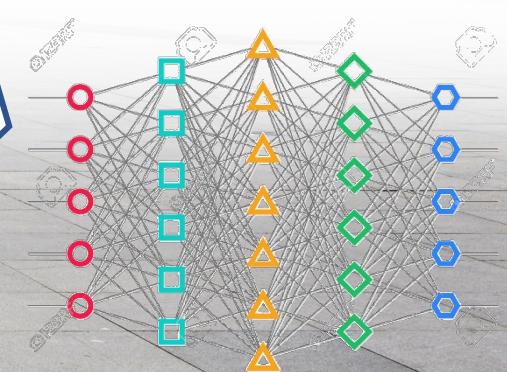
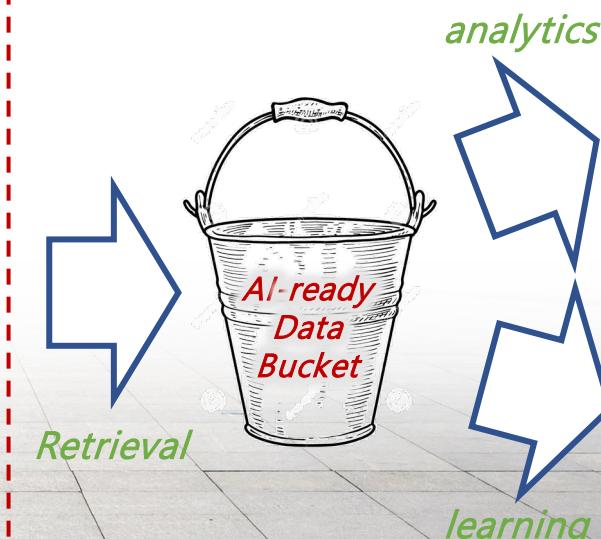
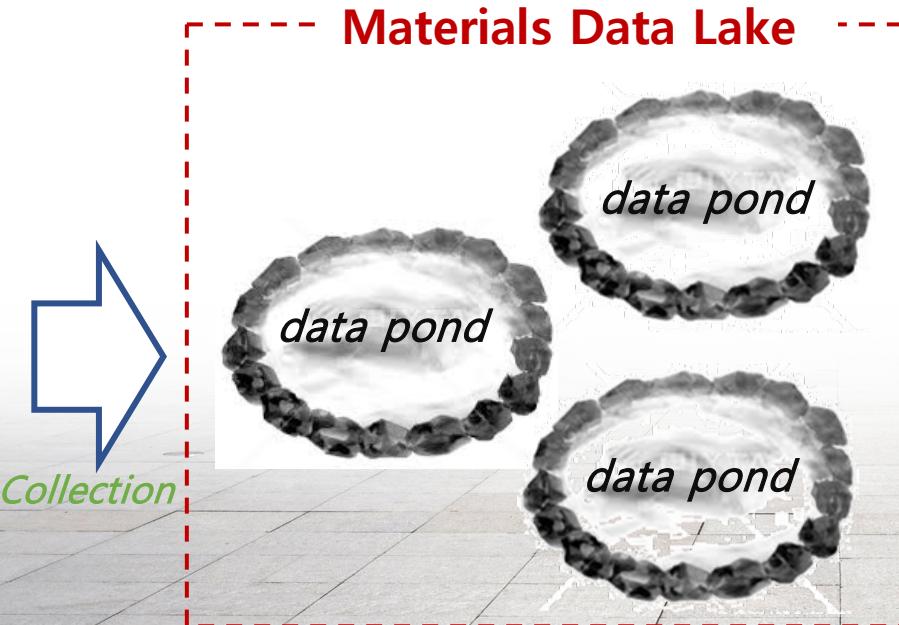
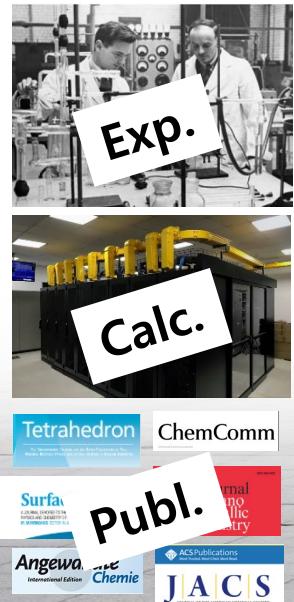


# AI-ready Data for Materials Design

- *Explicitly Defined Data*
- *Understandable Context of the Data* in the viewpoint of materials science



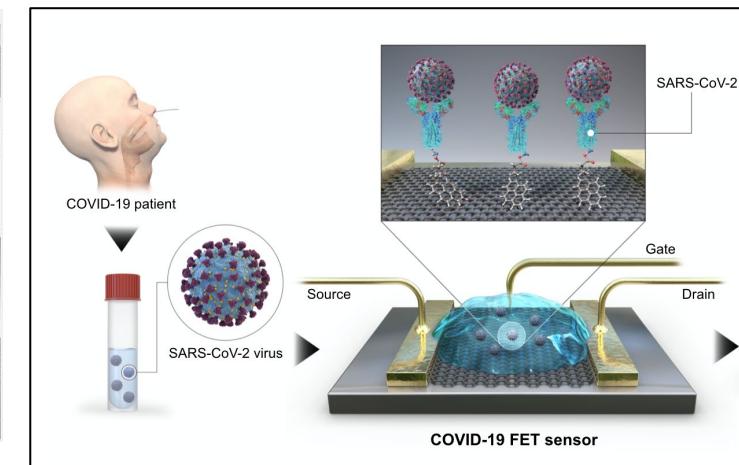
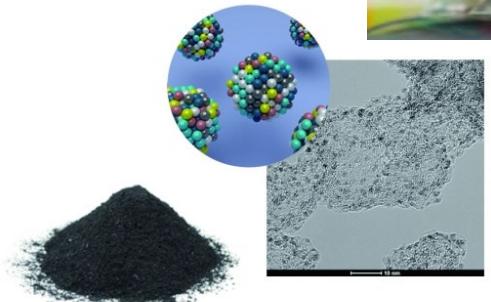
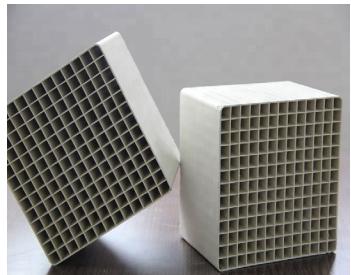
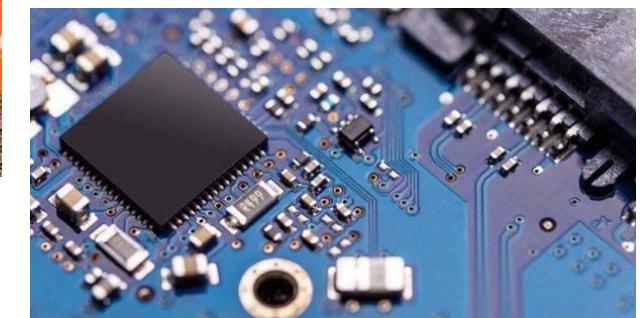
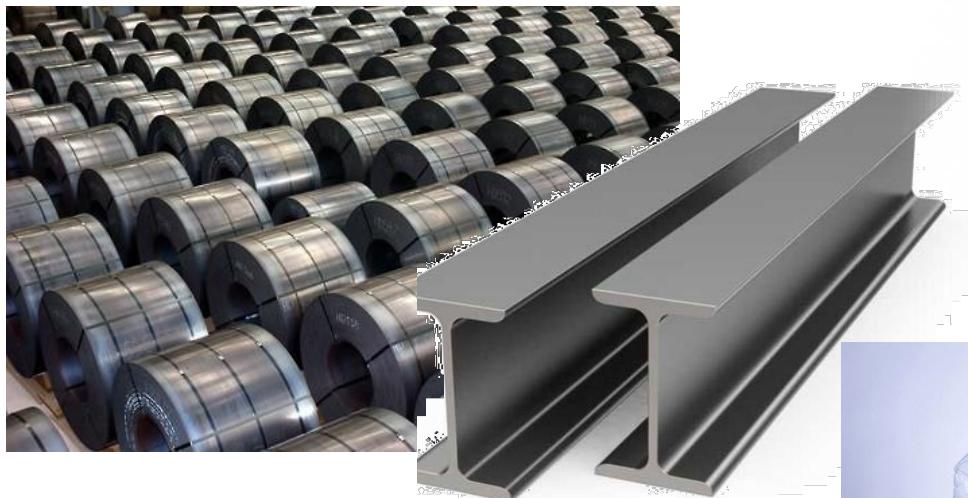
# Rock Base for the Data Driven Materials Design



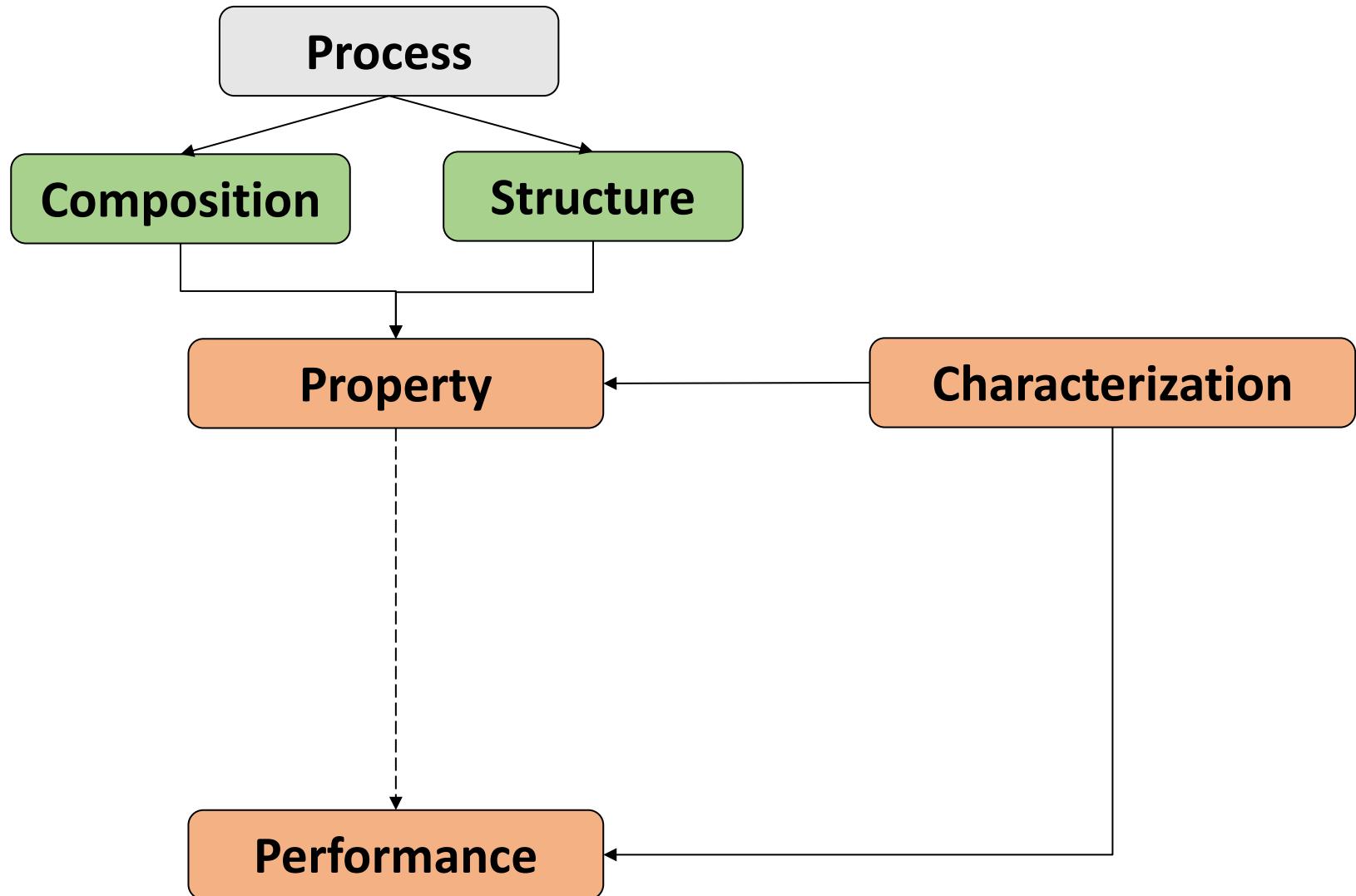
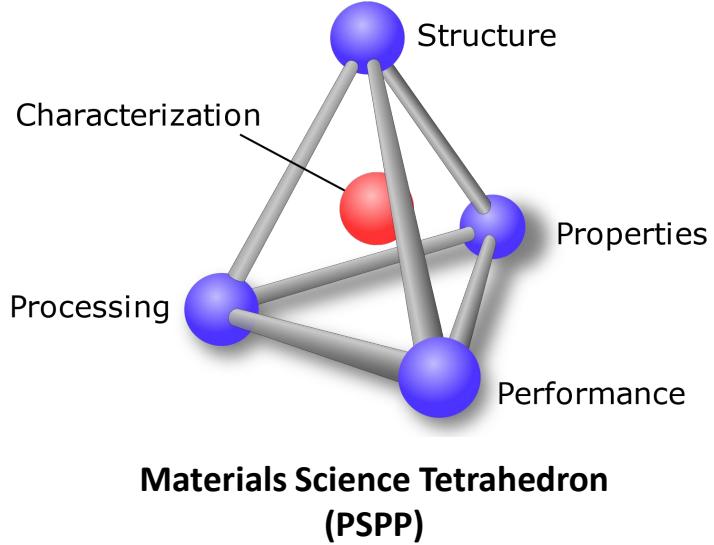
## Standard Data Schema and Vocabulary *Language for Research with Materials Data*

# Common Data Structure for Wide Spectrum of Materials?

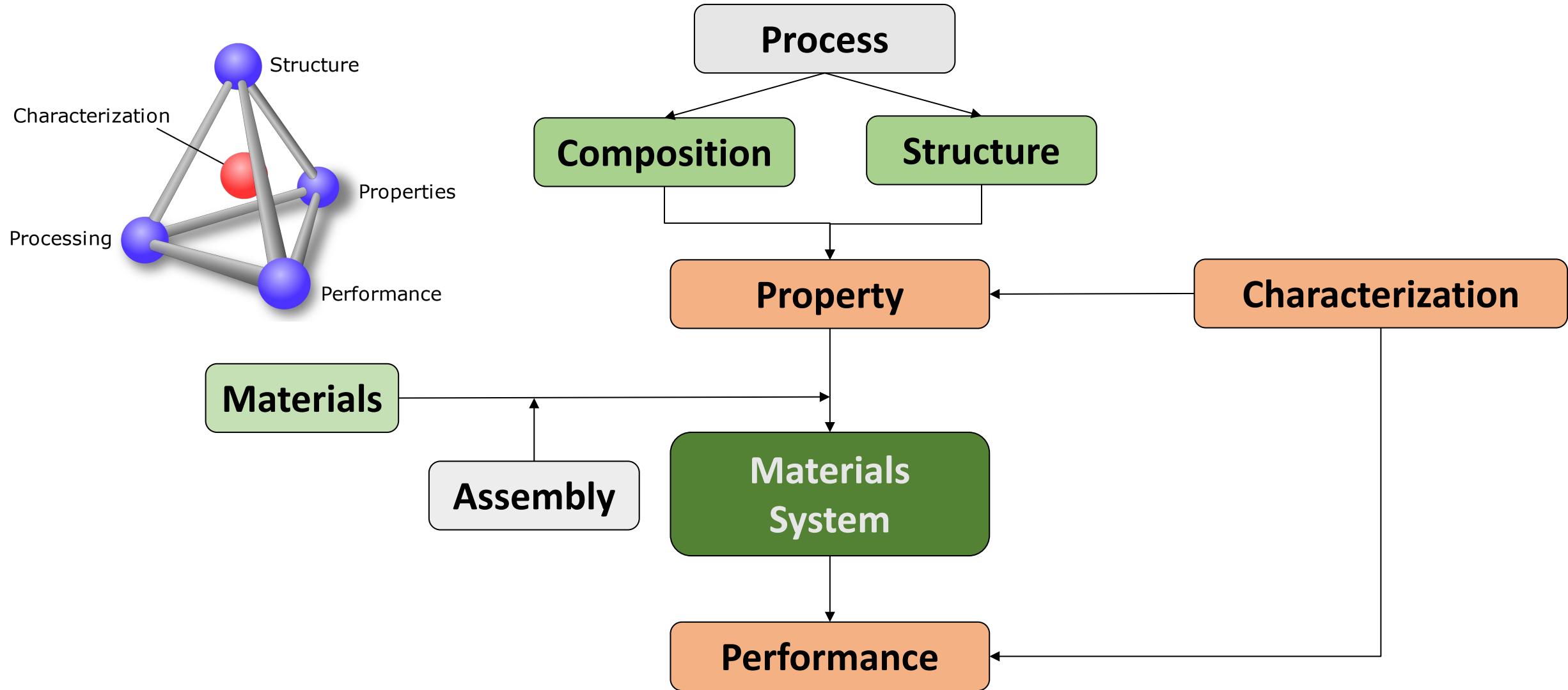
森羅萬象



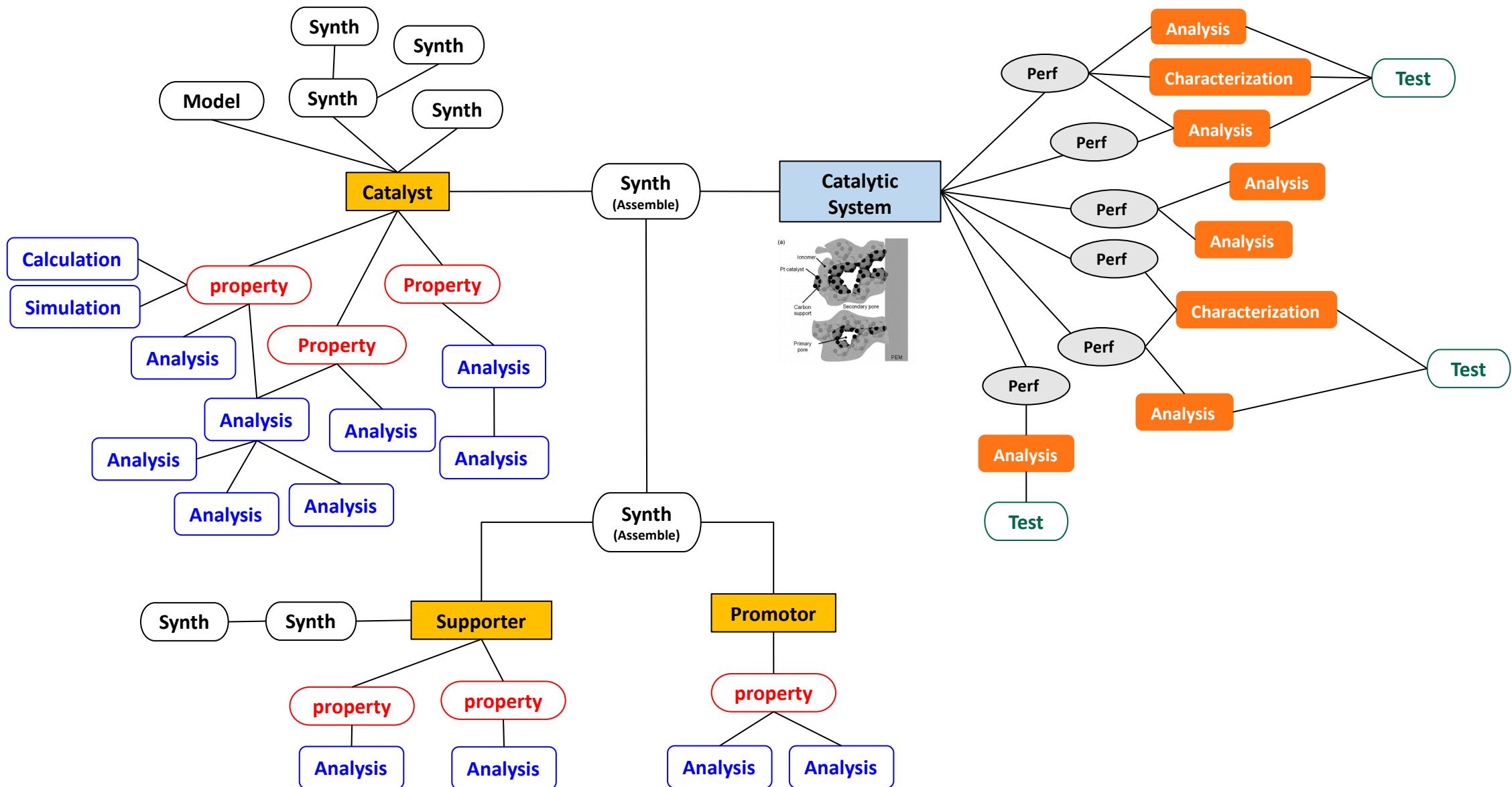
# Harmonized AI-ready Materials R&D Data



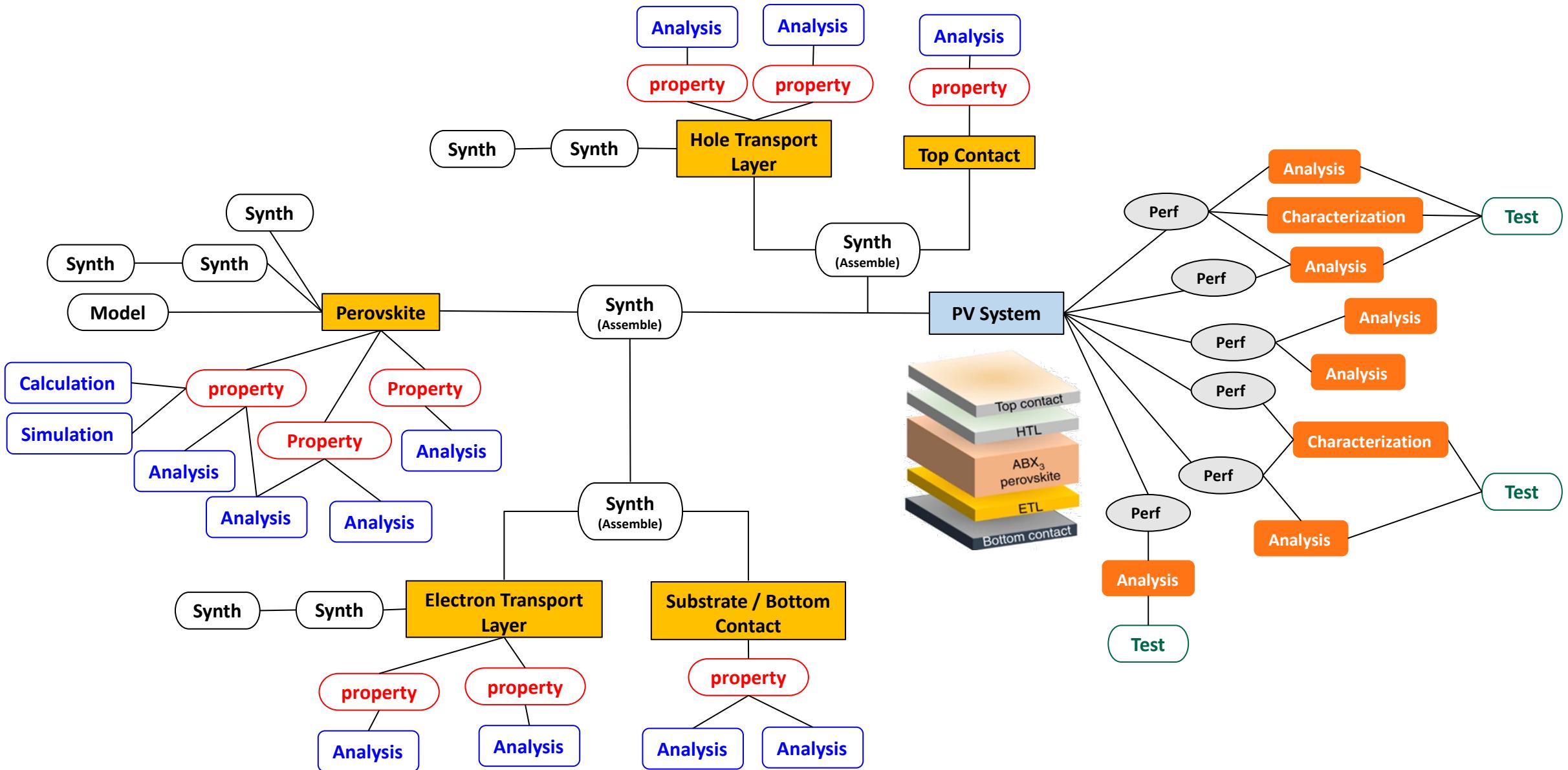
# Harmonized AI-ready Materials R&D Data



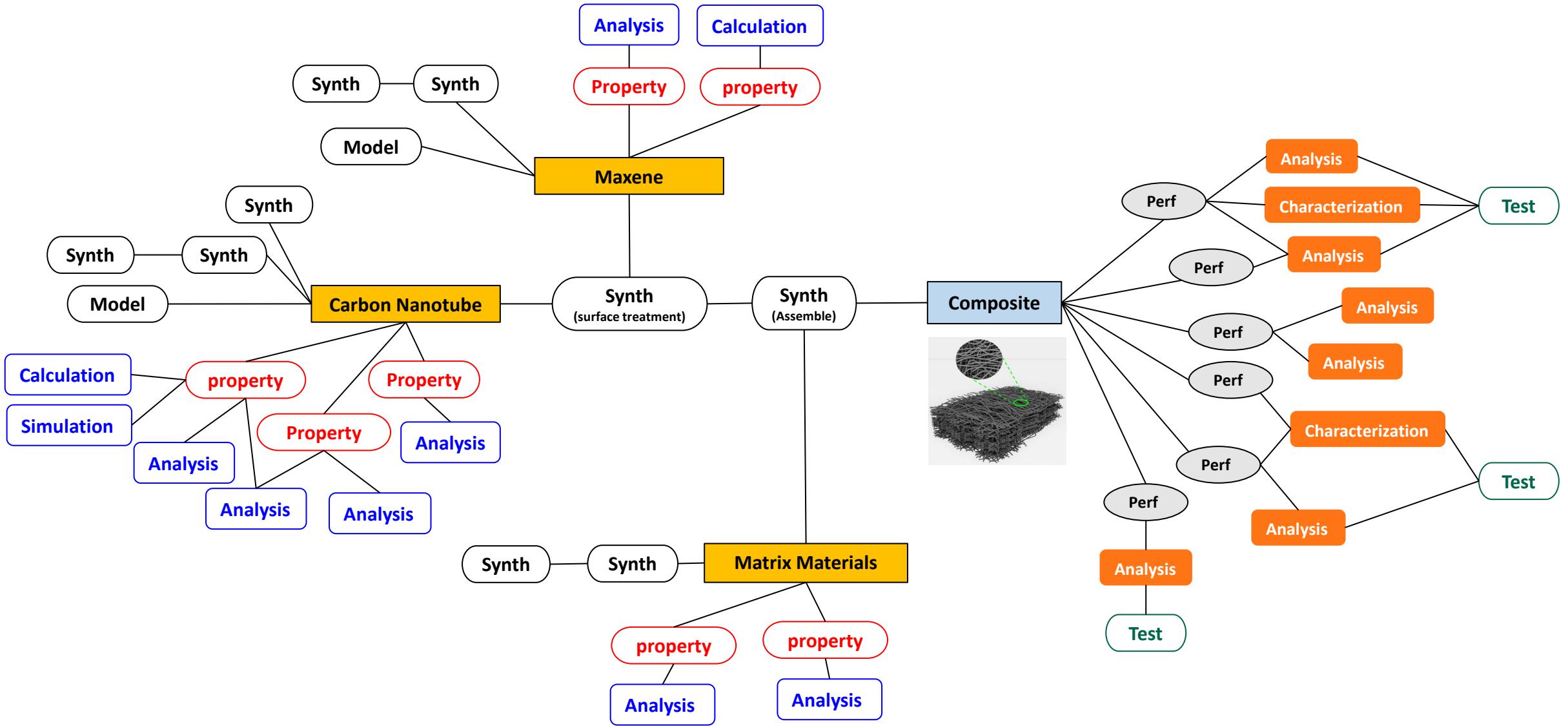
# Example 1 : Heterogeneous Catalysis



# Example 2 : Perovskite Solar Cell



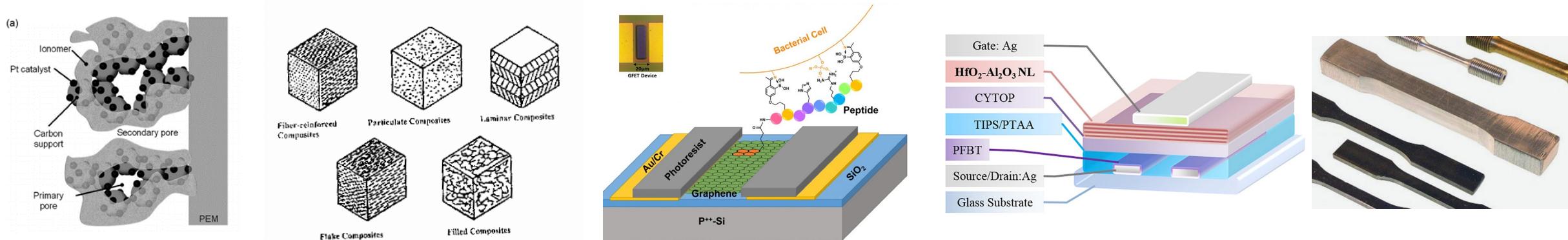
# Example 3 : Fiber Reinforced Materials



# Definition of Materials System

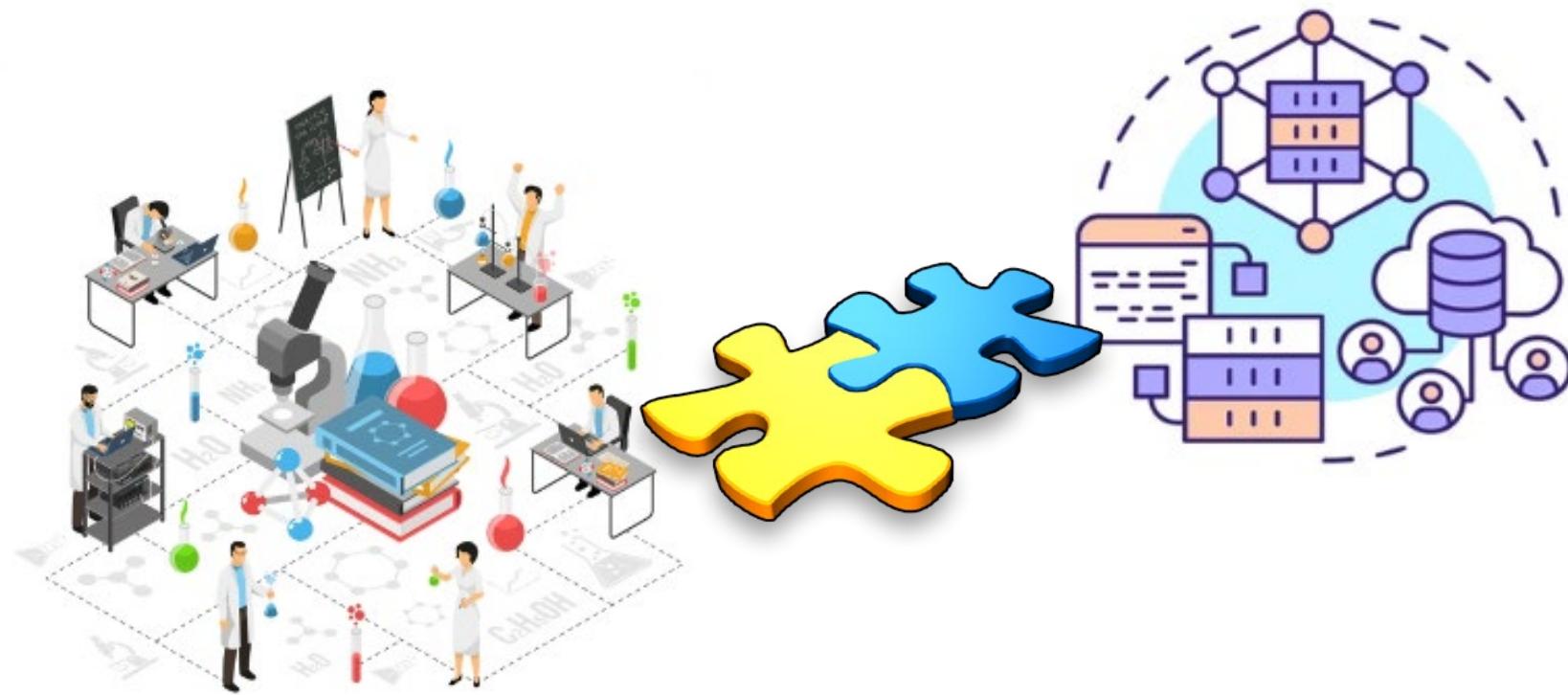
- Object that is composed of one or more materials to characterize the performance of a material.

Materials System Examples	Materials 1	Materials 2	Materials 3	Materials 4	...
Catalytic Materials	Catalyst	Supporter	Promotor		
Composite Materials	Fiber	Matrix			
Sensor Materials	Sensor	Electrode	Substrate	Protective Layer	
Cathode in Secondary Battery	Cathode 1	Cathode 2	Electrode	Electrolyte	
Chrome Coating Steel	Substrate Steel	Cr Coating Layer			
Special Steel	Special Steel				

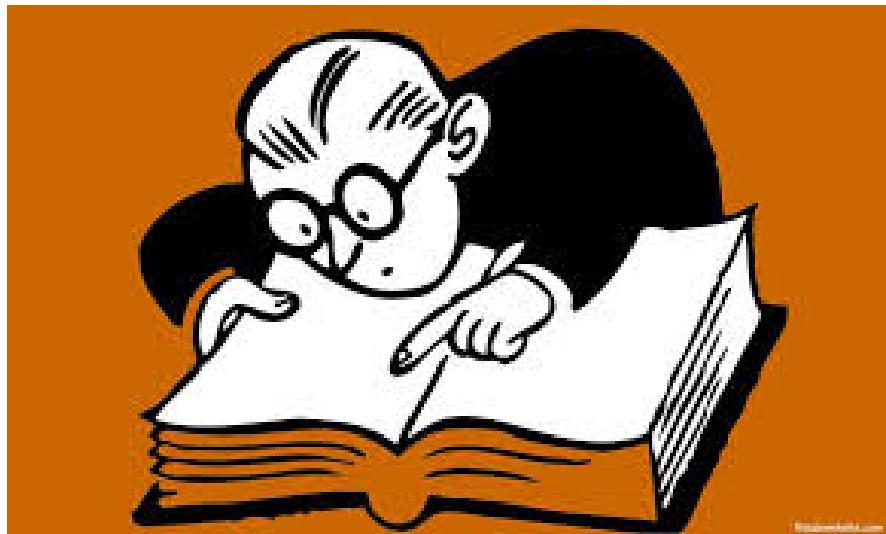


# Our Approach for Standardization

- We build the schema **based on the workflow of materials scientists.**
- We build the schema to cover whole sub-disciplines in materials science by **introducing the concept of “materials system”.**



# Outcome: Structured Data Keyword Dictionary



**Product Details**

Name:	Adjustable Race	List Price:	\$0.00
Product Number:	AR-5381	Standard Cost:	\$0.00
Model:	<input type="button" value="▼"/>	Sell Start Date:	1/06/1998 <input type="button" value="▼"/>
Category:	<input type="button" value="▼"/>	Sell End Date:	<d/MM/yyyy> <input type="button" value="▼"/>
Subcategory:	<input type="button" value="▼"/>	Discont. Date:	<d/MM/yyyy> <input type="button" value="▼"/>
Product Line:		Safety Stock:	1000
Class:		Reorder Point:	750
Style:		Days To Man.:	0
Color:		Make:	<input type="checkbox"/>
Size:	<input type="button" value="▼"/>	Finished Goods:	<input type="checkbox"/>
Weight:	<input type="button" value="▼"/>	OK Cancel	

**K-MDS**

About Data Search Support My Page

**Material innovation research  
To create an ecosystem**

**Korea Materials Data Station**

K-MDS is a material research data collection, sharing, and utilization platform established and operated at the national level.

We want to contribute to creating a materials data-based research and innovation ecosystem in Korea.

Materials Property Composition

Please enter your search term. (Use the special character)

Recommended keywords #Material #124912491

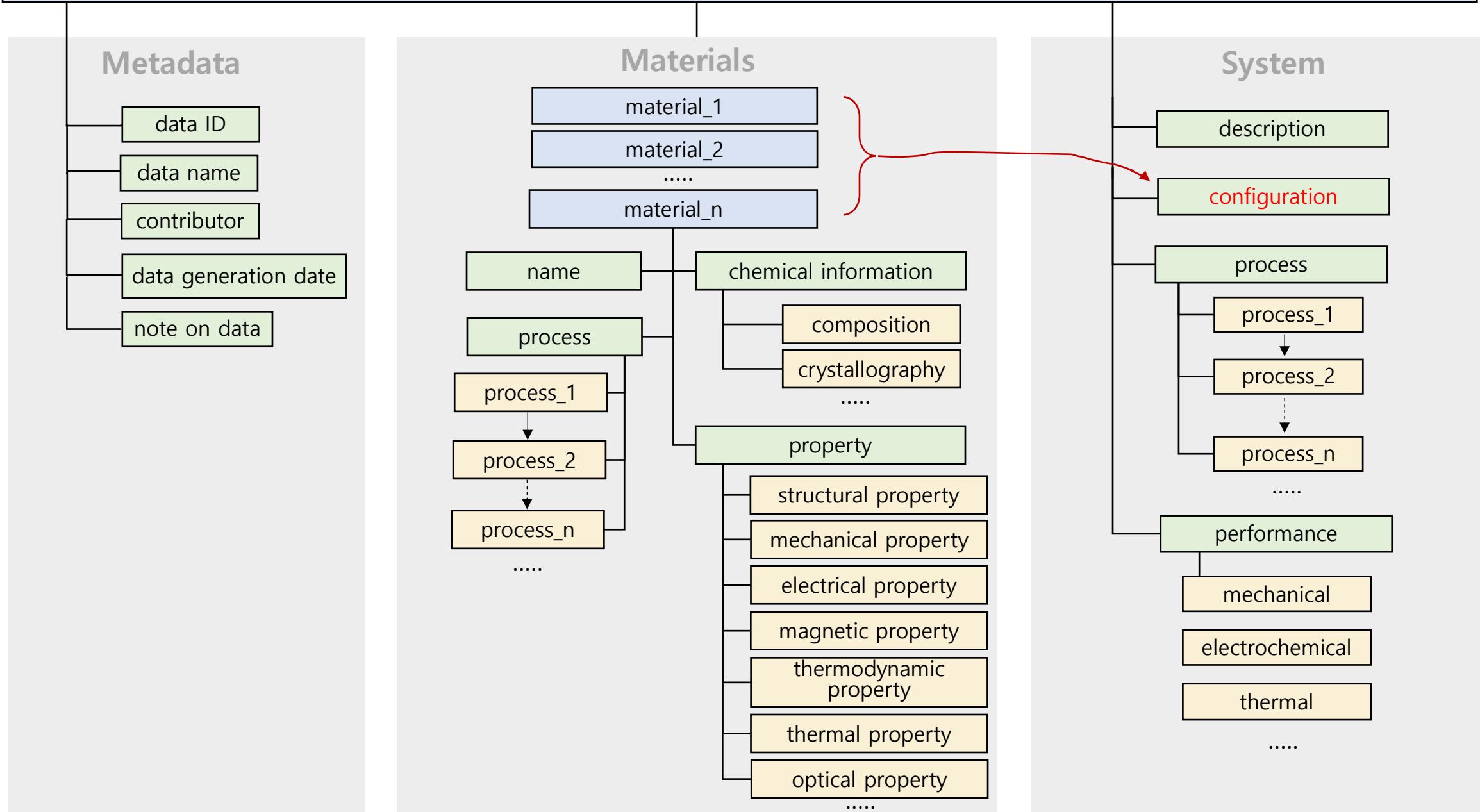
CATEGORY Data type

UTILIZATION Top search data Latest registration data Popular service

A circular network diagram centered around 'BandGap'. Other nodes include 'AI', 'Cr', 'DFT', 'Electrical', 'Mo', 'Optical', 'C', 'NI', 'TEM', and 'Thermal'. Lines connect nodes based on their properties, such as 'Electric' and 'Temperature'.

Korea Materials Data Station

# Materials Research Data Set



# Standard Structure of Materials Data Set

## metadata

- data ID
- data name
- contributor
- data generation date
- note on data

*Meta data section of the data set*

## materials\_n

- name
- chemical information
- process
- model
- property

*Data section of each n materials constructing the materials system  
Each materials' data is indexed by the system variable "Materials\_n"*

## system

- description
- configuration
- process
- performance

*Data section of Materials System*

```
▼ 0 {4}
  ▼ meta {5}
    data ID : DOI:111.222.333.444
    data     : Pt3Ni particle on graphite support data
    name    : 1
    ▶ contributor {2}
      data generation date : 2020-04-28
      note on  : information of material_2 comes from
      data     : Aldrich catalog\n10.1021/n1401881z
  ▼ material_1 {4}
    name : Pt3Ni
    ▶ chemical information {2}
    ▶ process [3]
      ▼ 0 {1}
        ▶ solvothermal {8}
      ▼ 1 {1}
        ▶ heat treatment {10}
      ▼ 2 {1}
        ▶ centrifugation {5}
    ▶ property {1}
  ▶ material_2 {2}
  ▼ system {4}
    description : Pt3Ni/C catalyst system for CO2 reduction
    ▶ configuration {3}
      active material : Material_1
      amount of active material : 0.003
      support material : Material_2
    ▶ process [3]
    ▶ performance {1}
      ▶ electrochemical {1}
        ▶ ORR {3}
          ▶ area-specific activity {3}
            value : 49
            uncertainty : 0.1
          ▶ measurement {1}
            ▶ electrochemical activity {5}
        ▶ mass-specific activity {3}
        ▶ Faradaic efficiency {3}
```

# Contents of the Structured Vocabulary Dictionary

- ▶ metadata vocabulary : { 8 props }
- ▶ materials vocabulary : { 1 prop }
- ▶ system specific vocabulary : { 7 props }
- ▶ common vocabulary for analysis method : { 47 props }
- ▶ common vocabulary for materials process : { 40 props }
- ▶ numeric data expression : { 2 props }

*3 vocabulary groups*  
*2 common vocabulary groups*  
*1 numeric data expression rule*

MatResData-Standard-Committee in [github.com](https://github.com)

어휘군	category 1	category 2	category 3
metadata vocabulary	data name		
	contributor	name	
		affiliation	
	data generation date		
materials common vocabulary	note on data		
		name	
	chemical information	composition	
		crystallography	
		SMILES	
	process	(process..n)	
		elements	
		state	
	model	box dimension	
		periodic boundary condition	
property		structure file	
		structural property	
		mechanical property	
		electrical property	
		magnetic property	
		thermodynamic property	
		thermal property	
		optical property	
		corrosion property	
system specific vocabulary	description		
	system_catalyst	active material	
		amount of active material	
		promotor	
	system_porous_materials	amount of promotor	
		support materials	
	process	(process..n)	
	system_memristi	electrochemical	
		thermal	
	performance	gas adsorption	
	description		

# Keyword format in the dictionary (in json format)

```
{"keyword": {"eng_definition": "string",
             "alias": "[string;..]",
             "data_type": "string (python)",
             "data_unit": "string",
             "data_example": "[string;..]"}
}
```

```
▼ name : {
    eng_definition : description of materials according to customary usage, designation
                     system, composition, supplier, brand name, et.
    alias : value
    data_type : string
    data_example : annealed INCONEL625 plate; Al7075-T6; Al10Si5Mg; Fe; CNT; TiO2; Ag; In202;
                  LLZO; Aldrich CFO plate
}
```

```
{"name": "Aldrich CFO plate"}
```

# Numeric Data Examples

```
▼ thermophysical property :{
    eng_definition : physical properties of materials affected by temperature, related to the structure
                    and dynamics of atoms
    alias : value
    ▼ thermal conductivity :{
        eng_definition : quantity of heat that passes in unit time through unit area of a substance
        alias : value
    ▼ value :{
        eng_definition : data value
        data_type : numeric
        data_unit : W m^{-1} K^{-1} (watts per meter-kelvin)
        data_example : 12.3
    }
    ▼ uncertainty :{
        eng_definition : uncertainty of the value
        data_type : numeric
        data_unit : same as value
    }
    ▼ measurement :{
        eng_definition : mesurement method
        alias : characterization tools, analysis method
        ▶ (analysis_n) :{ 2 props }
        ▶ sample description :{ 2 props }
    }
}
▶ thermal diffusivity :{ 5 props }
▶ thermal expansion coefficient :{ 5 props }
▶ activity :{ 9 props }
▶ boiling temperature :{ 5 props }
```

```
▼ thermophysical property :{
    ▼ thermal conductivity :{
        value : 12.3
        uncertainty : 0.02
    ▼ measurement :{
        ▼ thermalgravimetric analysis :{
            instrument : Perkin Elmer Series 3
            image : output-3.jpg
        }
    }
}
```

# Array Numeric Data

```
▼ numeric array data ("data_type" : "numeric array" :{
    format : Array data is composed of an array of list data. Each list in the
            array is correspondent to each row of data except 0th list in the
            array. 0th list is the column names. The data in the following list is
            to be aligned with the sequence of the column names of the 0th list.
            Each column is either independent (controlled) variables or dependent
            (measured) variables. "array_example" is to be provided for users to
            understand better the array data structure.
    ▼ column_definition :{
        eng_definition : column names to be listed in 0th list are defined in this
                          object.
    ▼ (column_name)) :{
        eng_definition : Definition of column i.e. name of column; it can be either
                          dependent or independent variables; uncertainty of the
                          value can be added in the following column; columns can be
                          defined as many as required for the data; listed in the
                          same sequence of the array data.
        data_type : (data type of the value in the column of array)
        data_unit : (data unit of the value in the column of array)
        data_example : (example of the value in the column of array)
    }
    }
    ▼ array_example : [ 5 items
        0 : ["column_name1","column_name2","column_name3"...]
        1 : [data of column1, data of column2, data of column3...]
        2 : [data of column1, data of column2, data of column3...]
        3 : [data of column1, data of column2, data of column3...]
        4 : [data of column1, data of column2, data of column3...]
    ]
}
```

Temperature	T <sub>uncertainty</sub>	Value	Value <sub>uncertainty</sub>
130	0.5	0.8	0.001
180	0.5	1.3	0.001
240	0.5	2.2	0.001

```
"array data example": [
    [ "temperature (K)",
        "uncertainty in temperature (K)",
        "yield stress (GPa)",
        "uncertainty in yield stress GPa)],
    [ 130, 0.5, 0.8, 0.001 ],
    [ 180, 0.5, 1.3, 0.001 ],
    [ 240, 0.5, 2.2, 0.001 ]
]
```

# Allowed Units based on 7 basic SI units

*red = 7 SI base units*

Time		Length		Mass	
year	year	m	meter	kg	kilogram
d	day	cm	centi meter	g	gram
h	hour	mm	milli meter	mg	milli gram
min	minute	{Wmu m}	micro meter	m_{e}	electron mass
s	second	nm	nano meter	Da	dalton
ms	milli second	WANGSTROM	Angstrom		
{Wmu s}	micro second				
ns	nano second				
fs	femto second				
Electricity		Thermodynamic Temperature / Energy / Power		Amount	
A	Ampere	eV	electron volt	mol	mole
mA	milli Ampere	keV	kilo electron volt		
{Wmu A}	micro Ampere	meV	milli electron volt		
nA	nano Ampere	J	Joule		
pA	pico Ampere	kJ	kilo Joule		
V	volt	pJ	pico Joule		
kV	kilo volt	W	watt		
mV	millivolt	K	kelvin degree		
{Wmu V}	micro volt	kW	kilo watt		
C	coulomb	MeV	mega electron volt		
{Wmu C}	micro Coulomb	mW	milli watt		
pC	pico Coulomb	{Wmu W}	micro watt		
e	charge of an electron				
S	conductance (Siemens)				
Ohm	resistance				
F	Faraday				

Luminous Intensity		Force / Pressure		Others	
cd	candela	N	Newton	at.%	atomic per cent
		{Wmu N}	micro Newton	wt.%	weight per cent
		kN	kilo Newton	ppm	part per million
		kgf	kilo gram force	%	per cent
		gf	gram force	HR	
		Pa	pascal	HB	
		MPa	mega pascal	HV	
		GPa	giga pascal	HK	
		Torr	Torr	Jones	specific detectivity
		bar	Bar	Hz	Hertz
		atm	atmosphere	kHz	kilo Hertz
				revolution	revolution, rotation
				degree	geometric angle
				L	liter
				mL	milli liter
				cycle	number of repeat
				T	tesla
				{Wmu L}	micro liter
				GHz	giga Hertz

소재 연구데이터 표준어휘 사전 작성  
가이드

Guide for Building the Standard Vocabulary  
Dictionary of Materials Research Data

2024.11.1.

소재 연구데이터 표준화 전문위원회

# Materials Vocabulary (as of 2024.10, 12 groups of property)

```
▼ materials vocabulary :{  
  ▼ material_(n) :{  
    eng_definition : materials that can be a component of materials system  
    alias : value  
    ▶ name :{ 4 props }  
  ▼ chemical information :{  
    eng_definition : atomic scale information of materials  
    alias : Stirring time  
    ▶ composition :{ 4 props }  
    ▶ purity :{ 5 props }  
    ▶ density :{ 5 props }  
    ▶ CAS number :{ 4 props }  
    ▶ chemical formula :{ 4 props }  
    ▶ structural formula :{ 4 props }  
    ▶ SMILES :{ 3 props }  
    ▶ molecular weight :{ 4 props }  
    ▶ additive :{ 4 props }  
  }  
  ▶ model :{ 7 props }  
  ▶ structure :{ 14 props }  
  ▼ property :{  
    eng_definition : attribute, quality, or characteristic of materials  
    alias : value  
    ▶ corrosion property :{ 18 props }  
    ▶ electrical property :{ 34 props }  
    ▶ electrochemical property :{ 3 props }  
    ▶ ferroelectric property :{ 5 props }  
    ▶ gas transmission property :{ 4 props }  
    ▶ magnetic property :{ 36 props }  
    ▶ mechanical property :{ 12 props }  
    ▶ optical property :{ 10 props }  
    ▶ piezoelectric property :{ 7 props }  
    ▶ rheological property :{ 5 props }  
    ▶ thermophysical property :{ 24 props }  
    ▶ tribological property :{ 3 props }  
  }  
  ▶ process :{ 3 props }  
}
```

```
▼ electrical property :{  
  eng_definition : physical properties that a material exhibits upon the application  
  of electric forces  
  alias : value  
  ▶ band gap bowing parameter :{ 5 props }  
  ▶ band gap energy :{ 5 props }  
  ▶ band gap type :{ 3 props }  
  ▶ band gap graded :{ 3 props }  
  ▶ carrier concentration :{ 5 props }  
  ▶ carrier diffusion length :{ 5 props }  
  ▶ carrier lifetime :{ 5 props }  
  ▶ conduction band minimum :{ 5 props }  
  ▶ charge of constituents :{ 5 props }  
  ▶ critical temperature :{ 5 props }  
  ▶ dielectric constant :{ 5 props }  
  ▶ dielectric strength :{ 5 props }  
  ▶ electric susceptibility :{ 5 props }  
  ▶ electrical conductivity :{ 5 props }  
  ▶ electronic density of states :{ 6 props }  
  ▶ electron effective mass :{ 5 props }  
  ▶ electron mobility :{ 5 props }  
  ▶ exciton binding energy :{ 5 props }  
  ▶ Fermi level :{ 5 props }  
  ▶ Hall coefficient :{ 5 props }  
  ▶ Hall voltage :{ 5 props }  
  ▶ hole effective mass :{ 5 props }  
  ▶ hole mobility :{ 5 props }  
  ▶ ionic conductivity :{ 5 props }  
  ▶ ionic mobility :{ 5 props }  
  ▶ ionic activation energy :{ 5 props }  
  ▶ permittivity :{ 5 props }  
  ▶ resistivity :{ 5 props }  
  ▶ valence band maximum :{ 5 props }  
  ▶ spin-Hall angle :{ 5 props }  
  ▶ spin-orbit splitting energy :{ 5 props }  
  ▶ temperature coefficient of resistance :{ 5 props }  
}
```

# System Specific Vocabulary *(as of 2024.10, 10 materials systems)*

```
▼ system specific vocabulary :{
  ▶ system_catalyst :{ [ 6 props ] }
  ▶ system_porous_materials :{ [ 6 props ] }
  ▶ system_memristive :{ [ 6 props ] }
  ▶ system_organic_thin_film_transistor :{ [ 6 props ] }
  ▶ system_photodetector :{ [ 6 props ] }
  ▶ system_gas_sensor :{ [ 6 props ] } ●
  ▶ system_alkali_ion_battery :{ [ 6 props ] }
  ▶ system_piezoelectric :{ [ 5 props ] }
  ▶ system_multilayer :{ [ 6 props ] }
  ▶ system_photovoltaic :{ [ 6 props ] }
}
```

```
▼ system_gas_sensor :{
  eng_definition : a device that detects the presence of gases
  alias : chemoresistive sensor; chemical sensor
  ▶ description :{ [ 3 props ] }
  ▼ configuration :{
    eng_definition : configurational information of the gas sensor
    ▶ substrate :{ [ 4 props ] }
    ▶ electrode :{ [ 7 props ] }
    ▶ channel :{ [ 5 props ] }
    ▶ interlayer :{ [ 5 props ] }
  }
  ▶ process :{ [ 2 props ] }
  ▼ performance :{
    eng_definition : performance data of the gas sensor
    ▶ measurement condition :{ [ 9 props ] }
    ▶ responsivity :{ [ 4 props ] }
    ▶ recovery time :{ [ 4 props ] }
    ▶ response time :{ [ 4 props ] }
    ▶ theoretical detection limit :{ [ 4 props ] }
    ▶ selectivity :{ [ 3 props ] }
    ▶ air stability :{ [ 3 props ] }
    ▶ humidity stability :{ [ 4 props ] }
    ▶ measurement :{ [ 4 props ] }
  }
}
```

# Process Data Construction (as of 2024.10, 53 unit processes)

```
▼ common vocabulary for materials process :{
  ▶ atomic layer deposition :{ 13 props }
  ▶ atomic layer etching :{ 13 props }
  ▶ ball milling :{ 16 props }
  ▶ blade coating :{ 7 props }
  ▶ bulk metal forming :{ 10 props }
  ▶ casting :{ 9 props }
  ▶ centrifugation :{ 8 props }
  ▶ chemical mechanical polishing :{ 12 props }
  ▶ chemical bath deposition :{ 7 props }
  ▶ chemical synthesis :{ 7 props }
  ▶ chemical vapor deposition :{ 11 props }
  ▶ diamond turning :{ 5 props }
  ▶ drying :{ 6 props }
  ▶ dialysis :{ 7 props }
  ▶ e-beam lithography :{ 11 props }
  ▶ electrochemical deposition :{ 13 props }
  ▶ electrospinning :{ 11 props }
  ▶ electrical poling :{ 8 props }
  ▶ encapsulation :{ 5 props }
  ▶ etching :{ 4 props }
  ▶ cleaning :{ 4 props }
  ▶ heat treatment :{ 8 props } ●
  ▶ hydrothermal reaction :{ 9 props }
  ▶ ion intercalation :{ 6 props }
  ▶ ink jet printing :{ 6 props }
  ▶ metal additive manufacturing :{ 22 props }
  ▶ microwave-assisted synthesis :{ 10 props } ●
  ▶ mixing :{ 7 props }
  ▶ molecular beam epitaxy :{ 11 props }
  ▶ perovskite postprocess :{ 4 props }
  ▶ polishing :{ 12 props }
  ▶ pressing :{ 5 props }
  ▶ pulsed laser deposition :{ 12 props }
  ▶ quenching :{ 4 props }
  ▶ rinsing :{ 5 props }
  ▶ sintering :{ 7 props }
  ▶ slurry casting :{ 8 props }
  ▶ slot coating :{ 4 props }
  ▶ solvent annealing :{ 5 props }
```

```
▼ heat treatment :{
  eng_definition : thermal process for the purpose of improving the properties of a
                    material, removing impurities(volatile substances), and/or incur thermal
                    decomposition
  alias : value
  ▶ method :{ 4 props }
  ▶ atmosphere :{ 4 props }
  ▶ temperature :{ 4 props }
  ▶ heating rate :{ 4 props }
  ▶ time :{ 4 props }
  ▶ cooling rate :{ 4 props }
}
```

```
▼ microwave-assisted synthesis :{
  eng_definition : microwave-assisted synthesis conditions that can be controlled during
                    process
  ▶ precursor_(n) :{ 3 props }
  ▶ solution_(n) :{ 3 props }
  ▶ heating rate :{ 5 props }
  ▶ temperature :{ 4 props }
  ▶ cooling rate :{ 5 props }
  ▶ atmosphere :{ 3 props }
  ▶ time :{ 4 props }
  ▶ microwave power :{ 5 props }
  ▶ microwave frequency :{ 5 props }
}
```

# Analysis Data Construction (as of 2024.10, 63 analysis methods)

```
▼ common vocabulary for analysis method :{
  ▶ AES : { 4 props }
  ▶ AFM : { 4 props }
  ▶ amperometry : { 6 props }
  ▶ atom probe tomography : { 4 props }
  ▶ compressive test : { 8 props }
  ▶ creep test : { 4 props }
  ▶ d33 meter : { 5 props }
  ▶ Dynamic Light Scattering(DLS) : { 5 props }
  ▶ DFT : { 22 props } ●
  ▶ DMA : { 4 props }
  ▶ differential scanning calorimetry (DSC) : { 4 props }
  ▶ EDS : { 5 props }
  ▶ EIS : { 8 props }
  ▶ electrochemical activity : { 7 props }
  ▶ empirical MD : { 6 props }
  ▶ EPMA : { 4 props }
  ▶ ERD : { 4 props }
  ▶ EQE analysis : { 5 props }
  ▶ EXAFS : { 4 props }
  ▶ fatigue test : { 13 props }
  ▶ film adhesion test : { 5 props }
  ▶ focussed ion beam : { 5 props }
  ▶ fsLA-ICP-MS : { 5 props }
  ▶ gas adsorption/desorption isotherm : { 8 props }
  ▶ gas chromatography : { 10 props }
  ▶ gel electrophoresis : { 6 props }
  ▶ liquid chromatography : { 4 props }
  ▶ hardness test : { 5 props }
  ▶ ICP-MS : { 5 props }
  ▶ ICP-OES : { 5 props }
  ▶ impact test : { 4 props }
  ▶ impedance analyzer : { 4 props }
  ▶ infrared spectroscopy : { 6 props }
  ▶ IV measurement : { 6 props }
  ▶ JV measurement : { 16 props } ●
  ▶ magnetic hysteresis loop : { 7 props }
  ▶ memristive activity : { 8 props }
  ▶ nano-indentation : { 4 props }
  ▶ NEXAFS : { 4 props }
  ▶ nuclear magnetic resonance : { 5 props }
```

```
▼ DFT :{
  eng_definition : quantum mechanical calculations using density functional theory
  ▶ code : { 4 props }
  ▶ calculation mode : { 4 props }
  ▶ basis : { 3 props }
  ▶ charge : { 5 props }
  ▶ energy cutoff : { 5 props }
  ▶ optimizer : { 4 props }
  ▶ exchange-correlation functional : { 4 props }
  ▶ solvent model : { 3 props }
  ▶ k-point : { 3 props }
  ▶ convergence criteria : { 4 props }
  ▶ symmetry : { 4 props }
  ▶ magnetic structure : { 4 props }
  ▶ potential : { 4 props }
  ▶ LDA+U : { 6 props }
  ▶ partial occupation : { 4 props }
  ▶ total energy : { 5 props }
  ▶ density of state : { 6 props }
  ▶ phonon calculations : { 7 props }
  ▶ nudged elastic band : { 7 props }
  ▶ ab-initio MD : { 7 props }
  ▶ spin orbit coupling : { 4 props }
}
```

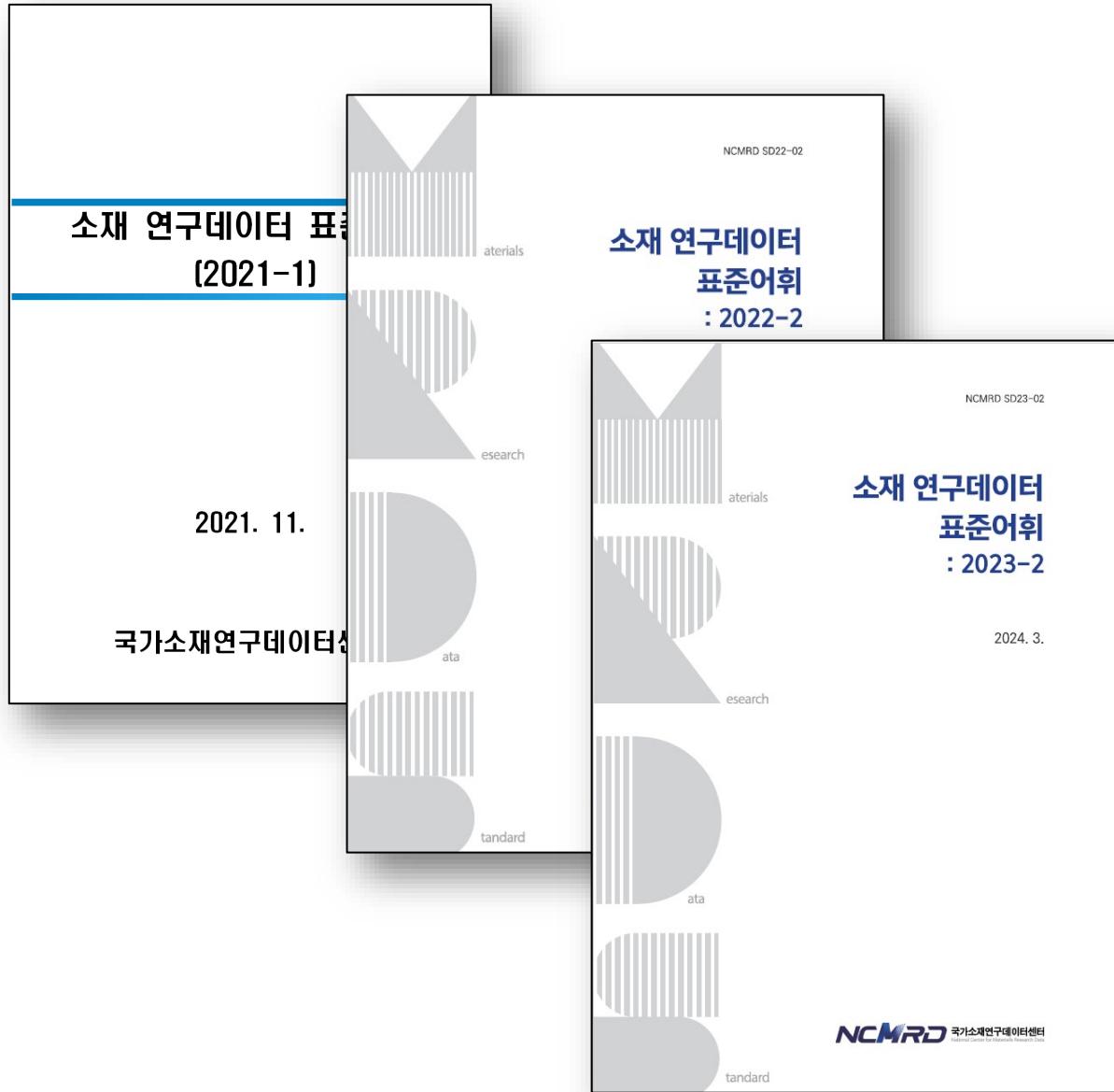
```
▼ JV measurement :{
  eng_definition : IV measurement of solar cell as it is illuminated
  alias : value
  ▶ equipment : { 3 props }
  ▶ stabilized test : { 5 props }
  ▶ number of cells averaged : { 5 props }
  ▶ certified : { 4 props }
  ▶ certification institute : { 4 props }
  ▶ storage before measurement : { 5 props }
  ▶ atmosphere : { 4 props }
  ▶ relative humidity : { 5 props }
  ▶ temperature : { 5 props }
  ▶ light source : { 5 props }
  ▶ light : { 8 props }
  ▶ scan : { 8 props }
  ▶ preconditioning : { 6 props }
  ▶ raw data : { 4 props }
}
```

# MatResData-Standard-Committee in github.com

The screenshot shows the GitHub repository page for 'MatResData-Standard-Committee'. The repository is public and was generated from [krlee227/first-trial](#). The main branch is not protected. The repository contains 6 branches and 1 tag. The 'Code' tab is selected. The repository has 32 commits. Key files include 'README.md', '.gitattributes', 'Paper 재료마당 Jaeryo Madang (2022).pdf', 'Presentation @ NIST 221014.pdf', 'README.md', and 'Rules for Dictionary 240119.pdf'. The 'About' section describes the repository as a working space for Materials Research Data Standard Expert Committee, mentioning kmds.re.kr/en/. The 'Branches' section lists branches: main, 2021-1, 2022-1, 2022-2, 2023-1, and 2023-2. The 'Tags' section is empty. The 'Releases' section shows a single release: 'dictionary 2022-2 released [Latest] on Dec 6, 2022'. The 'Packages' section is empty.

The screenshot shows the GitHub interface with a modal open for switching branches. The modal title is 'Switch branches/tags' and it contains a search bar with 'Find or create a branch...'. Below the search bar is a table with two tabs: 'Branches' (selected) and 'Tags'. It lists branches: main (default), 2021-1, 2022-1, 2022-2, 2023-1, and 2023-2. A checkmark is next to 2023-2. At the bottom of the modal is a button 'View all branches'. The background shows the main repository page for 'MatResData-Standard-Committee'. The repository has 6 branches and 1 tag. The 'Code' tab is selected. The repository has 49 commits. Key files include 'Initial commit', 'Update README.md', and 'Add files via upload'. The 'About' section describes the repository as a working space for Materials Research Data Standard Expert Committee, mentioning kmds.re.kr/en/. The 'Branches' section lists branches: main, 2021-1, 2022-1, 2022-2, 2023-1, and 2023-2. The 'Tags' section is empty. The 'Releases' section shows a single release: 'dictionary 2022-2 released [Latest] on Dec 6, 2022'. The 'Packages' section is empty.

# Standard Vocabulary Dictionary Publication



정 보 마 당

### 소재 연구데이터의 구조 및 표준 어휘 사전 구축<sup>1)</sup>

김수현<sup>1</sup>, 김정현<sup>2</sup>, 김지한<sup>3</sup>, 신호선<sup>4</sup>, 안재평<sup>5</sup>, 오창석<sup>6</sup>, 이광렬<sup>7\*</sup>, 이동화<sup>8</sup>, 이정훈<sup>9</sup>, 박선화<sup>4</sup>, 장현주<sup>9</sup>, 한상수<sup>7</sup>, 한승우<sup>10</sup>, 허용학<sup>11</sup> (가나다순)

<sup>1)</sup> 한국재료연구원 금속재료연구본부, <sup>2)</sup> 국립현대대학교 신소재공학과, <sup>3)</sup> 한국과학기술원 생명화학공학부, <sup>4)</sup> 한국표준과학연구원 스마트소시ety, <sup>5)</sup> 한국과학기술연구원 연구지원센터 디자인본부, <sup>6)</sup> 한국재료연구원 재료디자임플랫폼연구본부, <sup>7)</sup> 한국과학기술연구원 철단소재연구본부 계산과학연구센터, <sup>8)</sup> 포항공과대학교 신소재공학과, <sup>9)</sup> 한국화학연구원 화학데이터기반연구센터, <sup>10)</sup> 서울대학교 공과대학 재료공학부 및 신소재공동연구소, <sup>11)</sup> 국가소재연구데이터센터 한국표준과학연구원

\*E-mail: krlee@kist.re.kr

#### 초록

체계적으로 구조화된 소재 연구 데이터는 그 자체로 활용도가 높은 연구개발 자원일 뿐 아니라, 인공지능에 의한 소재 설계 예측을 갖추기 위한 핵심 연구기반이다. 다양한 소스로부터 수집된 소재 연구 데이터를 효율적으로 검색하고 활용하기 위해서는 데이터의 이름 등 키워드를 통일하고 해당 데이터의 태입 그리고 수치 데이터의 단위를 일치시키는 것이 대단히 중요하다. 또한, 미신라닝에 의한 소재 설계를 위해서는 소재의 성능 데이터를 소재의 구조나 공정상의 특이성과 연계하여 수집 관리하여야 한다. 이 상관관계를 기계적으로 학습함으로써 소재 설계의 미신라닝 모델이 만들어질 수 있기 때문이다. 따라서, 데이터 키워드의 표준화와 함께 데이터의 구조가 표준화되어야 소재 연구 데이터의 활용성을 크게 증진시킬 수 있을 것이다. 소재 연구 데이터 표준화 전문위원회에서는 다양한 소재 연구

#### I. 서론

현대 초연결 사회를 실현한 ICT 기술의 비약적 발전은 소재 분야에서도 연구개발 방법론 자체에 큰 변화를 일으키고 있다. 합성과 분석을 반복하는 실험적 연구와 열역학 및 kinetics에 바탕을 둔 이론적 해석 연구가 소재 개발의 고전적인 접근이었다면, ICT 기술의 혁신적 발전은 전자, 원자 수준의 계산과학을 이용한 소재 설계와 데이터에 기반한 소재 정보학 등 새로운 개념의 효율적인 연구개발 수단을 제공하고 있다. 방대한 양의 소재 데이터는 이를 체계적으로 분석하는 것만으로도 소재 탐색의 방향성을 제시할 수 있기 때문에 실험 연구의 효율성을 크게 향상시킬 수 있을 것으로 기대하고 있다. 최근에는 한 발짝 더 나아가 축적

<sup>1)</sup> 본 논문은 한국연구재단의 지원을 받는 국가 소재 연구데이터 센터, 소재 연구데이터 표준화 전문위원회의 연구 결과이다.

# Summary

---

- We have been building a harmonized data schema and vocabularies for whole sub-disciplines of materials science and technology.
- This schema is based on the workflow of materials research, easy to understand for materials researcher.
- Collaboration for mapping schema would be the first step for global scale (maybe) federated materials R&D database.
- We welcome your participation for building and extending vocabularies in this dictionary.

***We are fully open to any suggestions and volunteering for more useful language of material research in the future.***