Ontology Building

Facilitating Materials Ontology Development via NLP and Machine Learning

Starting with a corpus of 5893 PDFs

MS&T 2011, 2012

TMS 2009, 2010, 2011, 2012

NIST Diffusion Data Center

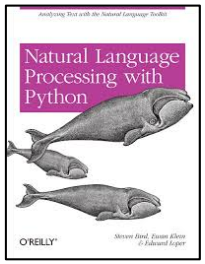
www.matscitech.org www.tms.org patapsco.nist.gov/diffusion

1 → 2 → 3 → 4

OCROPUS

thoonk.py

redis



gensim

python

radimrehurek.com/gensim

www.nltk.org

Focus of current development

5893 PDFs - 5.8 GB

63K pages

3.8M lines

53GB

3.4M lines extracted

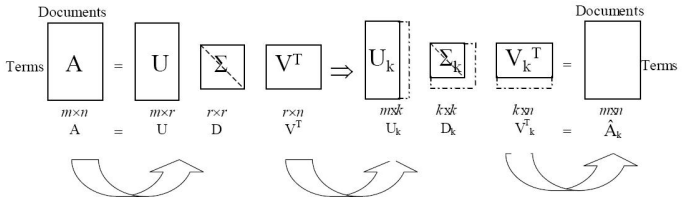
20M tokens

Extract text[1,2]

```
temperatur 69079
materi     53673
diffus     49743
process    48594
result     43713
surfac     33899
```

Identify key stems by frequency[3]

Use Latent Semantic Indexing to group similar concordance entries[4]



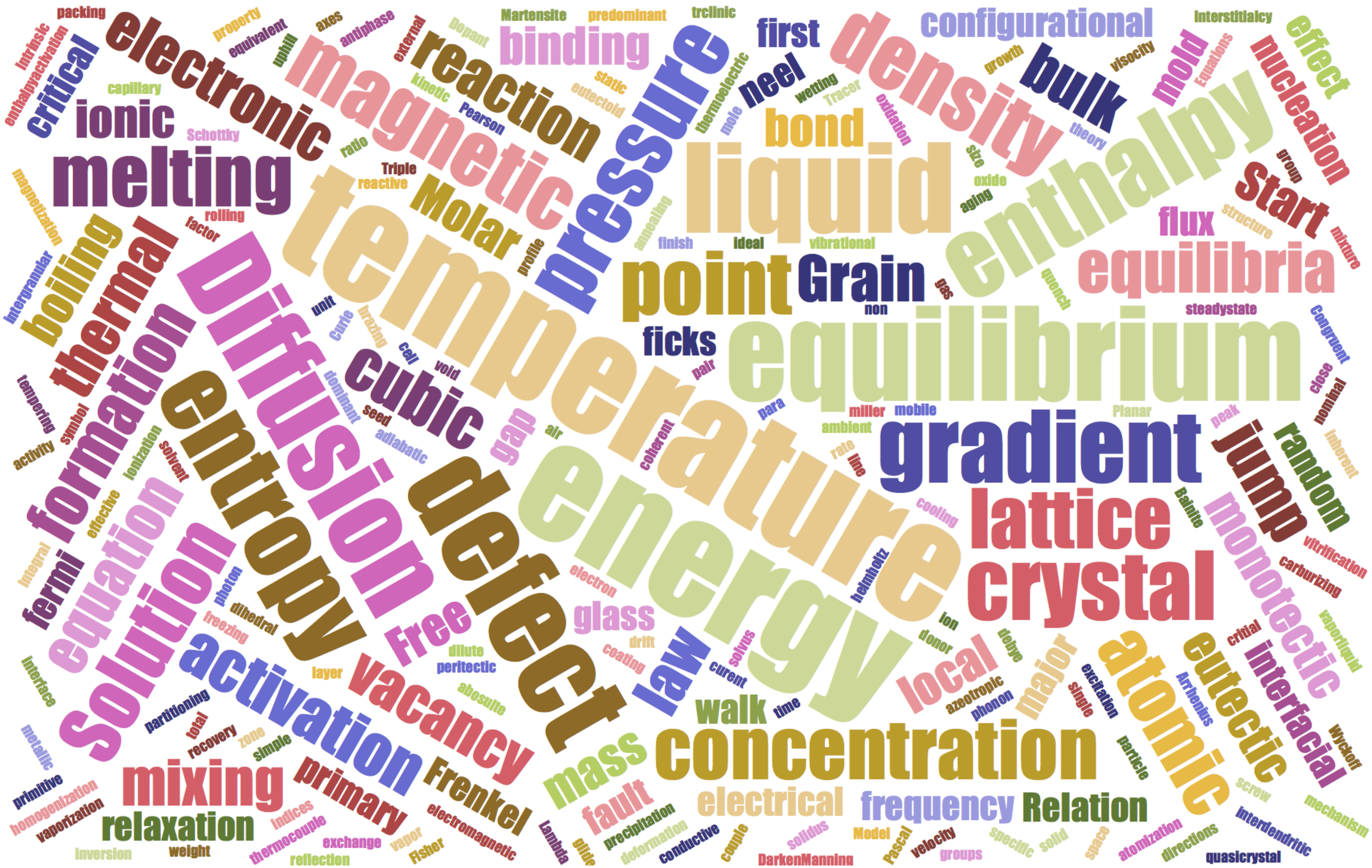
http://liqiangguo.wordpress.com/2011/06/09/latent-semantic-analysis/

Generate concordances for key stems[3]

```
tivity method. Lattice diffusion coefficients and grain boundary diffusion coe
nts and grain boundary diffusion coefficients were measured. Plots of the data
r all the alloys , the diffusion coefficients at tempe- ratures 1150 'C and hi
, were only accepted. From these coefficients , the parameters of temperature
 those alloys the self-diffusion coefficients of which were also determined at
ting intercrystallite diffu sion coefficients , the following rel ation coeffi
 place. Values of bulk diTfusion coefficients at temperatures below 1j50 ' C w
```

Alden Dima - 2/24/2014

TERMS


temperature
energy
equilibrium
liquid
enthalpy
Diffusion
defect
entropy
pressure
magnetic
density
gradient
concentration
crystal
lattice
point
melting
electronic
reaction
atomic
eutectic
monotectic
activation
vacancy
solution
equation
formation
thermal
boiling
cubic
mixing
relaxation
primary
Frenkel
mass
fault
electrical
frequency
Relation
walk
local
law
glass
Free
equilibria
Grain
ficks
bond
binding
bulk
first
neel
configurational
nucleation
effect
mold
Start
flux
jump
random
interfacial
major
Molar
ionic
critical
fermi
gap
packing
Martensite
predominant
triclinic
Interstitialcy
capillary
Schottky
vibrational
ideal
finish
steadystate
Congruent
miller
mobile
ambient
Planar
peak
cooling
vitrification
carburizing
helmholtz
electron
peritectic
dilute
layer
drift
coating
partitioning
recovery
simple
total
primitive
homogenization
vaporization
inversion
weight
thermocouple
indices
exchange
reflection
vapor
Fisher
electromagnetic
lambda
glide
precipitation
deformation
conductive
couple
DarkenManning
Model
Pascal
velocity
groups
specific
solid
space
atomization
directions
quasicrystal
interdendritic
mechanism
Wyckoff
particle
phonon
azeotropic
time
curent
solvus
solidus
donor
ion
debye
Single
excitation
Arrhenius
critial
screw
vaporliquid
close
nominal
inherent
mixture
Structure
group
Equations
viscosity
theory
growth
oxidation
Tracer
wetting
mole
thermoelectric
size
oxide
aging
quench
gas
non
eutectoid
static
kinetic
Pearson
Dopant
external
antiphase
axes
equivalent
uphill
property
intrinsic
enthalpyvacivation
ratio
Triple
reactive
rolling
factor
magnetization
annealing
profile
intergranular
unit
Curie
brazing
cell
void
pair
para
seed
adiabatic
dominant
coherent
rate
line
air
symbol
tempering
activity
ionization
solvent
effective
integral
photon
dihedral
freezing
interface
metallic
zone
absolute

Key Terms

Energy

Liquid

Concentration

Gradient

Temperature

Entropy

Diffusion

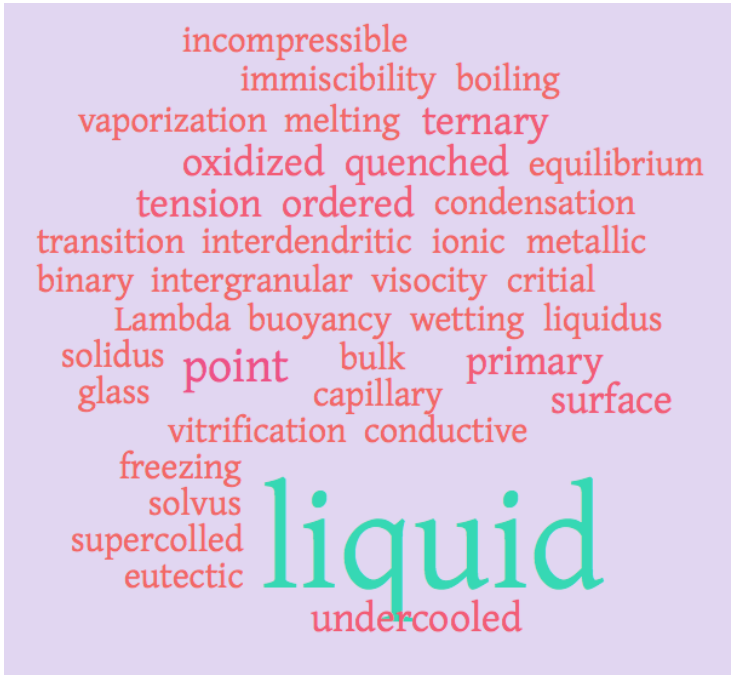
Defect

Crystal

Equilibrium

Enthalpy

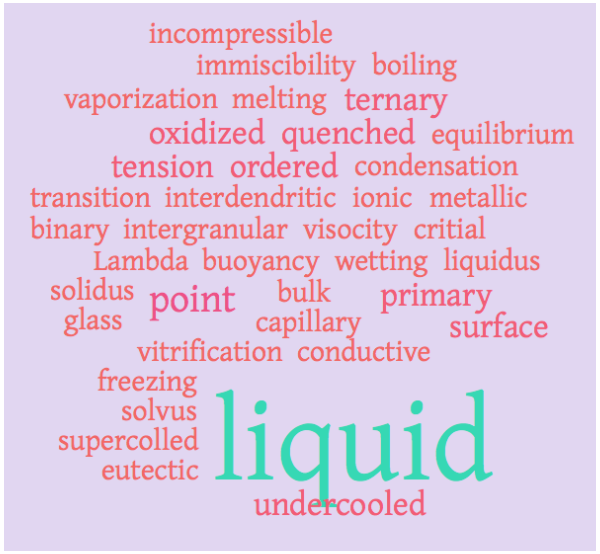
Validate terms



Cross-out any terms that do not belong.

binary liquid
bulk liquid
conductive liquid
equilibrium liquid
eutectic liquid
interdendritic liquid
intergranular liquid
ionic liquid
~~metallic liquid~~
ordered liquid
oxidized liquid
primary liquid
quenched liquid
ternary liquid
undercooled liquid
viscosity
boiling point
immiscibility.
condensation
critical point
freezing
Lambda point
~~solvus~~
solidus
liquidus
oxidized liquid
primary liquid
quenched liquid
ternary liquid
undercooled liquid
ordered liquid
surface tension
wetting
incompressible
capillary
surface tension
buoyancy
melting
glass transition
super cooled
vaporization
vitrification

Add Missing Terms



Add any missing relevant terms

binary liquid
bulk liquid
conductive liquid
equilibrium liquid
eutectic liquid
interdendritic liquid
intergranular liquid
ionic liquid
~~metallic liquid~~
ordered liquid
oxidized liquid
primary liquid
quenched liquid
ternary liquid
undercooled liquid
viscosity
boiling point
immiscibility.
Condensation
critical point
freezing
Lambda point
solidus
liquidus
oxidized liquid
primary liquid
quenched liquid
ternary liquid
undercooled liquid
ordered liquid
surface tension
wetting
incompressible
capillary
surface tension
buoyancy
melting
glass transition
super cooled
vaporization
vitrification

Terms added at workshop

congruent melting
miscibility gap
superheated
flash point (burn temperature)
solidification
mushy zone
dew point
density
heat capacity
ordering
pre-melting
incipient melting
convention
lamellar
vapor pressure
Scheil model
dendrite
segregation
color
slag
peritectic
monotectic
amorphous
azeotropic transition

Organize/Cluster Terms: Example for liquid terms

binary liquid
bulk liquid
liquidus
equilibrium liquid
eutectic liquid
interdendritic liquid
ionic liquid
~~metallic liquid~~
ordered liquid
oxidized liquid
primary liquid
quenched liquid
ternary liquid
undercooled liquid
viscosity
boiling point
immiscibility.
condensation
critical point
freezing
Lambda point
solidius
conductive liquid
oxidized liquid
primary liquid
intergranular liquid
undercooled liquid
ordered liquid
surface tension
wetting
incompressible
capillary
surface tension
buoyancy
melting
glass transition
super cooled
vaporization
vitrification

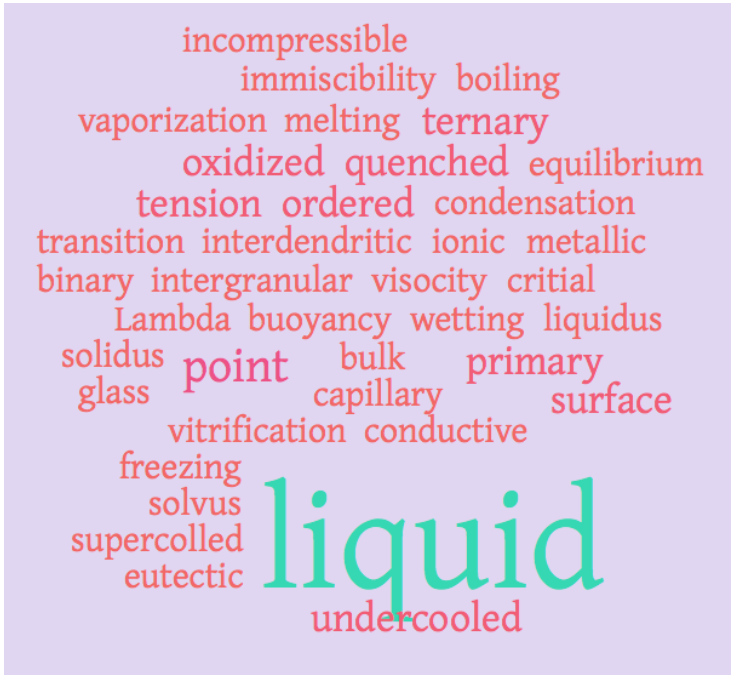
Binary liquid
Ternary liquid
Solvus
Liquidus
Critical point
Eutectic liquid

Freezing
Undercooled liquid
Quenched liquid
Interdendritic liquid
Primary liquid
Bulk liquid

Wetting
Capillary
Surface tension
Buoyancy

Organize/cluster terms;
Note relationships between terms

Topics Related to Terms



Subject areas associated with terms

Binary liquid
Ternary liquid
Solvus
Liquidus
Critical point
Eutectic liquid

PHASE DIAGRAM

Solidification/Casting

Freezing
Undercooled liquid
Quenched liquid
Interdendritic liquid
Primary liquid
Bulk liquid

Solder;Sintering

Wetting
Capillary
Surface tension
Buoyancy

Temperature TERMS

- activation temperature
- invariant temperature
- adiabatic temperature
- magnetic temperature
- aging temperature
- melting temperature
- air temperature
- metallization temperature
- ambient temperature
- mold temperature
- annealing temperature
- monotectic temperature
- azeotropic temperature
- boiling temperature
- nitriding temperature
- bond temperature
- brazing temperature
- nominal temperature
- carburizing temperature
- nucleation temperature
- coating temperature
- ordering temperature
- coherent temperature
- oxidation temperature

quench temperature
equilibrium temperature
reaction temperature
eutectic temperature
reactive temperature
eutectoid temperature
recovery temperature
external temperature
rolling temperature
homogenization temperature
tempering temperature
interfacial temperature
thermocouple temperature
Glass Transition Temperature
Liquidus Temperature
melting temperature
Congruent temperature
metallization temperature
Austenite Start Temperature
mold temperature
Austenite finish temperature
monotectic temperature
Bainite Start Temperature
neel temperature
Martensite Start Temperature
nitriding temperature
absolute temperature

Terms added at workshop

cooling temperature
partitioning temperature
critical temperature
peritectic temperature
Curie temperature
precipitation temperature
Debye temperature
vibrational energy
Einstein temperature
thermocouple
phase transformation
pyrometer
radiation
black body
recrystallization temperature
solution temperature
Celsius
Kelvin
Fahrenheit
DTA
DSC
liquidus
solvus
solidius
cooling rate
heating rate
Phonons

Crystal Terms

- Grain
- crystal growth
- crystal lattice
- crystal orientation
- crystal structure
- space group
- crystal defect
- polycrystal
- single crystal
- nanocrystalline
- miller indices
- planes
- directions
- cubic
- lattice system
- close packing
- bravais lattice

point groups
defects
impurity
unit cell
lattice parameter
d-spacing
face-centered cubic
body-centered cubic
simple cubic
trclinic
monoclinic
orthorhombic
rhombohedral
tetragonal
hexagonal
cubic

atomic coordination
quasicrystal
reflection
rotation
inversion
screw axes
glide planes
dislocation
Brillouin zone
liquid crystal
primitive crystal
seed crystal
Wyckoff position
Strukturbericht designation
Pearson symbol

Diffusion Terms

- Activation Energy
- Grain boundary diffusion
- Diffusion Coefficient
- Tracer diffusion
- Concentration profile
- Diffusion couple
- Diffusion layer
- Diffusion rate
- Solute Diffusion
- Impurity Diffusion
- Intrinsic Diffusion
- drift velocity
- Divacancies
- self-interstitials
- Frenkel Disorder
- diffusion flux
- relaxation time
- Darken-Manning Equations
- short-circuit
- substitutional
- surface diffusion
- uphill diffusion
- Dopant Diffusion
- Bulk diffusion
- Correlation factor
- diffusion distance
- Fick's first law
- thin-film solution
- random walk
- atomic jump
- jump frequency
- interdiffusion
- Kirkendall effect
- Boltzmann-Matano method
- Darken Equation
- Vacancy
- Interstitial
- steady-state diffusion
- vacancy diffusion
- point defects
- Schottky Disorder
- Nernst-Einstein Relation
- segregation
- Nernst-Planck Equation
- Arrhenius Relation
- Activation volume
- Interstitialcy mechanism
- Fisher model
- chemical diffusion

Defect Terms

- atomic defect
- bond defect
- dominant defect
- donor defect
- electronic defect
- frenkel defect
- equilibrium defect
- inherent defect
- lattice defect
- line defect
- local defect
- major defect

mobile defect
ordered defect
oxide defect
Planar defect
point defect
predominant defect
thermal defect
Triple defect
Vacancy defect
void defect

Energy Terms

- activation energy
- antiphase boundary energy
- atomic energy
- atomization energy
- binding energy
- bond energy
- boundary energy
- bulk energy
- chemical energy
- configurational energy
- critical energy
- defect energy
- deformation energy
- dihedral energy
- effective energy
- electrical energy
- electromagnetic energy
- electron energy

electronic energy
equivalent energy
exchange energy
excitation energy
fault energy
fermi energy
formation energy
free energy
gap energy
grain boundary energy
gradient energy
helmholtz energy
interfacial energy
ion energy
ionization energy

kinetic energy
lattice energy
magnetic energy
magnetization energy
migration energy
mixing energy
nucleation energy
ordering energy
pair energy
peak energy
phonon energy
photon energy
potential energy
reaction energy
relaxation energy
thermal energy
thermoelectric energy
vibrational energy

Enthalpy Terms

- Activation Enthalpy
- Binding Enthalpy
- Formation Enthalpy
- Free Enthalpy
- Integral Enthalpy
- magnetic enthalpy
- melting enthalpy
- migration enthalpy
- mixing enthalpy
- molar enthalpy
- partial enthalpy
- reaction enthalpy
- total enthalpy
-

Concentration Terms

- molality
- normality
- mole fraction
- mass fraction
- solution
- solute
- miscible
- saturated
- supersaturation
- phase separation
- suspension
- ideal solution
- solubility
- partition coefficient
- mixing ratio
- dilute solution
- solid solution
-
-

Entropy Terms

- activation entropy
- binding entropy
- boundary entropy
- configurational entropy
- electronic entropy
- formation entropy
- magnetic entropy
- melting entropy
- migration entropy
- mixing entropy
- molar entropy
- thermal entropy
- vibrational entropy
- fusion entropy

Gradient Terms

- activity gradient
- chemical gradient
- concentration gradient
- electrical gradient
- energy gradient
- potential gradient
- property gradient
- temperature gradient
- thermal gradient

Equilibrium Terms

- chemical equilibria
- invariant equilibria
- local equilibria
- thermodynamic equilibria
- univariant equilibria
- atomic equilibrium
- binary equilibrium
- chemical equilibrium
- defect equilibrium
- dynamic equilibrium
- eutectic equilibrium
- interface equilibrium
- invariant equilibrium
- ionic equilibrium
- liquid/liquid equilibrium
- local equilibrium
- monotectic equilibrium
-

monotectic equilibrium

non equilibrium

para equilibrium

reaction equilibrium

ternary equilibrium

thermal equilibrium

thermodynamic equilibrium

univariant equilibrium

vacancy equilibrium

vapor/liquid equilibrium

equilibrium phase