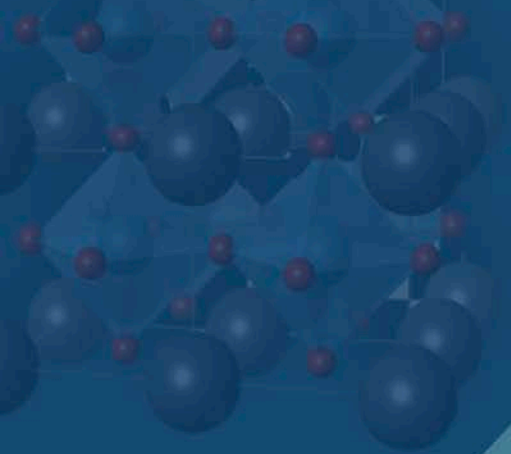
Materials Design and Discovery Using Learning Machines

An Example of Scintillator Discovery

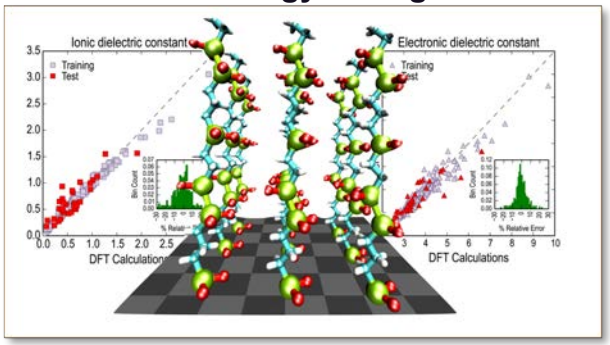
Ghanshyam Pilania
Materials Science and Technology Division
Los Alamos National Laboratory



11:00 AM (MST) Monday August 10, 2020

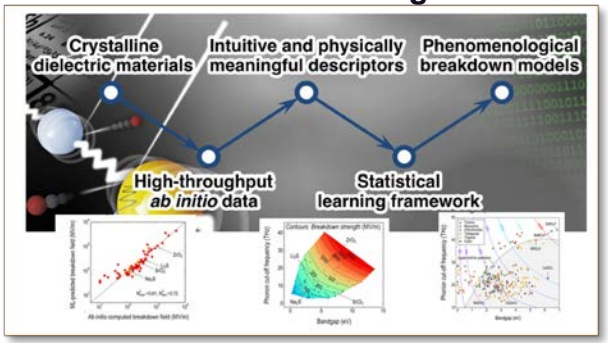
Computational Design of Materials

Designing Polymer Dielectrics for Energy Storage



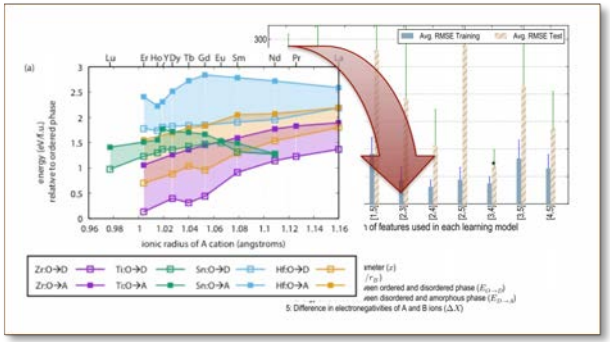
Sci. Rep. 6, 20952 (2016) & **Comput. Mater. Sci.** 125 123 (2016).

Learning Models for Dielectric Breakdown Strength



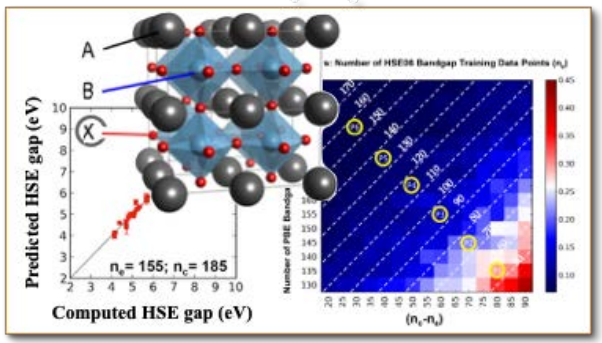
Chem. Mater. 28, 1304 (2016) & **J. Phys. Chem. C** 120, 14575 (2016)

Understanding Radiation Damage Resistance



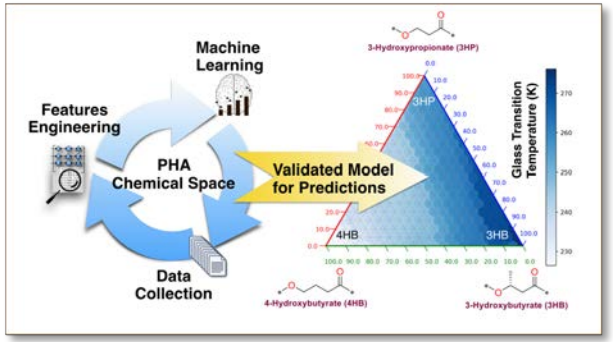
Chem. Mater. 29, 2574 (2017)

Learning Bandgaps Solids



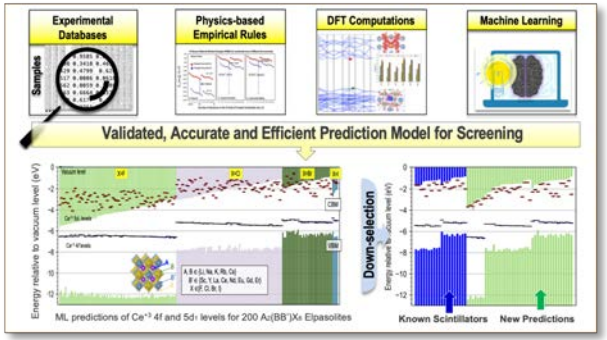
Sci. Rep. 6 19375 (2016) & **Comput. Mater. Sci.** 129 156 (2017).

Biodegradable Polymer Design



J. Chem. Inf. Model. 59, 5013 (2019)).

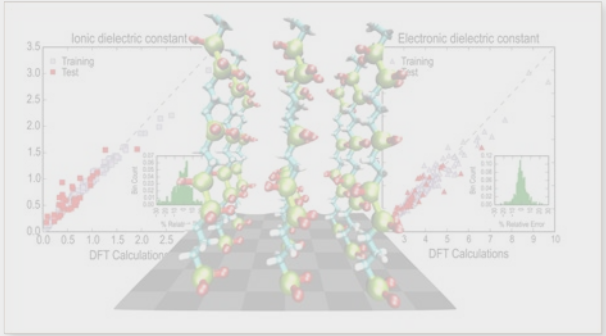
Scintillator Discovery



J. Chem. Phys. 148 241729 (2018).

Focus of This Talk: Scintillators Discovery

Designing Polymer Dielectrics for Energy Storage

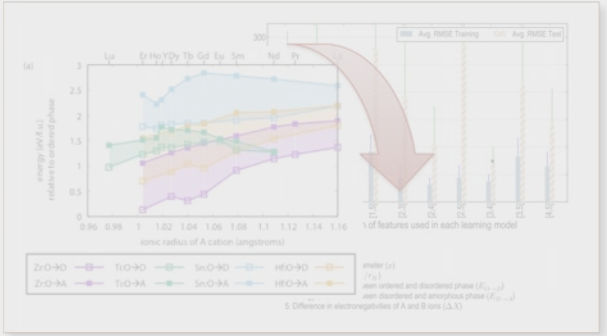


Sci. Rep. 6, 20952 (2016) & **Comput. Mater. Sci.** 125 123 (2016).

Learning Models for Dielectric Breakdown Strength

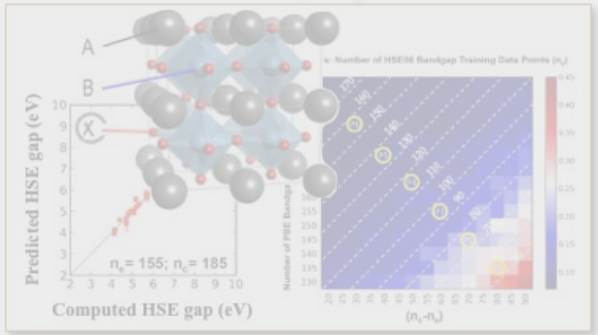
Chem. Mater. 28, 1304 (2016) & **J. Phys. Chem. C** 120, 14575 (2016)

Understanding Radiation Damage Resistance



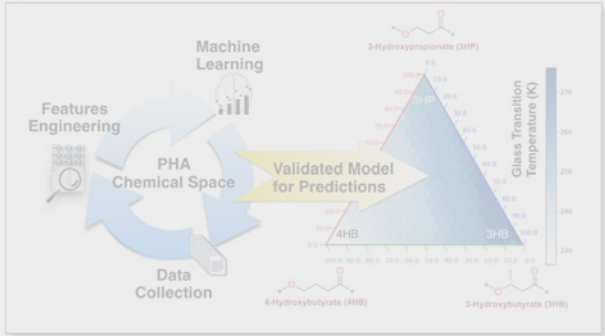
Chem. Mater. 29, 2574 (2017)

Learning Bandgaps Solids



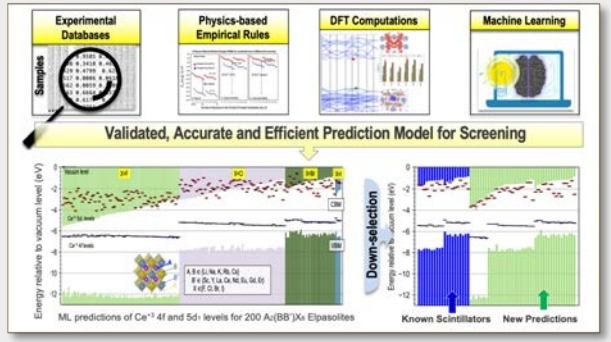
Sci. Rep. 6 19375 (2016) & **Comput. Mater. Sci.** 129 156 (2017).

Biodegradable Polymer Design



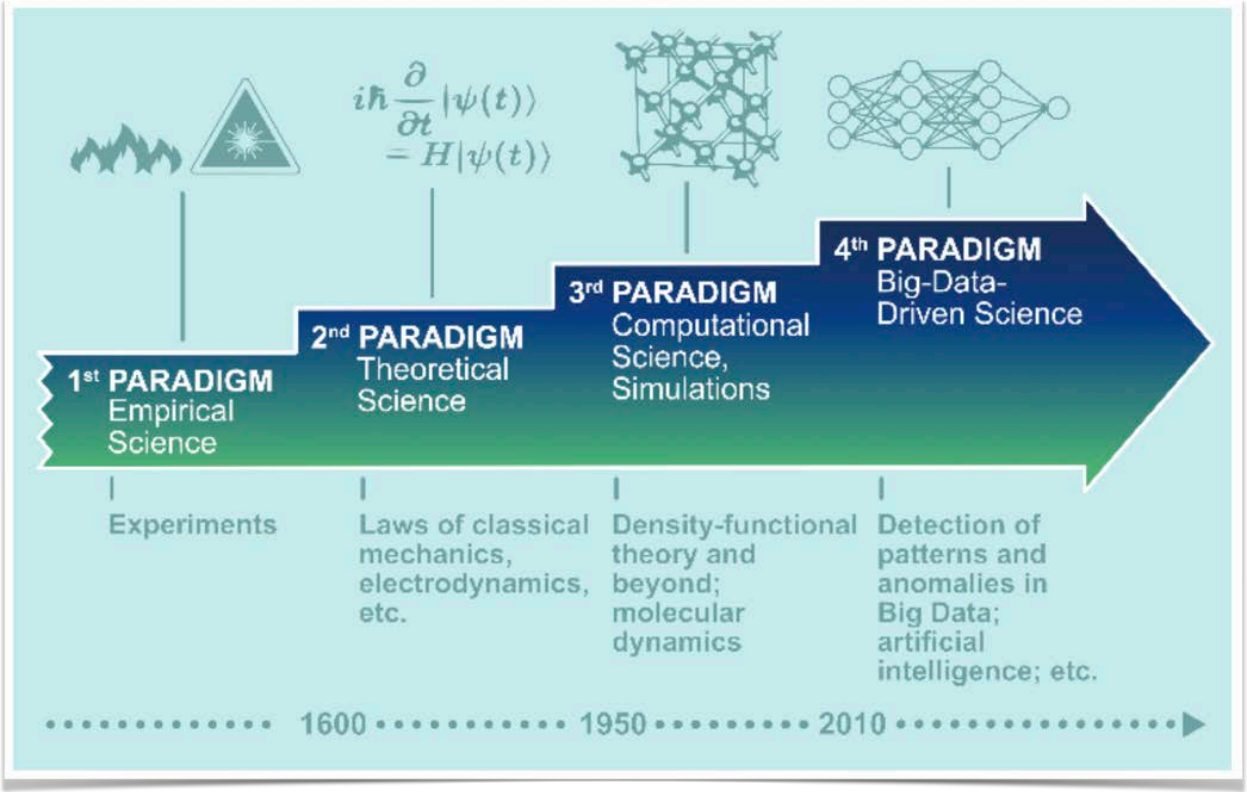
J. Chem. Inf. Model. 59, 5013 (2019)).

Scintillator Discovery



J. Chem. Phys. 148 241729 (2018).

Machine Learning in Materials Science



Recent progress reviewed in:

- G. Pilania, P. V. Balachandran, J. E. Gubernatis, T. Lookman T. *Data-Based Methods for Materials Design and Discovery: Basic Ideas and General Methods.* Synthesis Lectures on Materials and Optics. (2020).
- J. Schmidt, M. R. Marques, S. Botti, M. A. Marques, *Recent advances and applications of machine learning in solid-state materials science*. NPJ Computational Materials. 8, 1 (2019).
- R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, *Machine learning in materials informatics: Recent applications and prospects.* NPJ Computational Materials. 13, 1 (2017).

C. Draxl, M. Scheffler, *NOMAD: The FAIR Concept for Big-Data-Driven Materials Science*, MRS Bulletin 43 676 (2018).

Open and FAIR Materials Data

Open-Source Materials Databases

The Materials Project

NOMAD

Khazana 1.0
Polymer Genome Data Repository Graphical Visualization

OpenKIM

NIST
STANDARD REFERENCE DATA

AFLOW
Automatic-FLOW for Materials Discovery

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

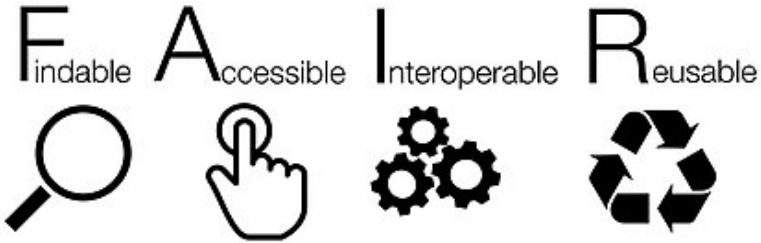
High Performance Computing Center
Materials Database

OQMD

Citrine

SCIENTIFIC DATA

and many more…

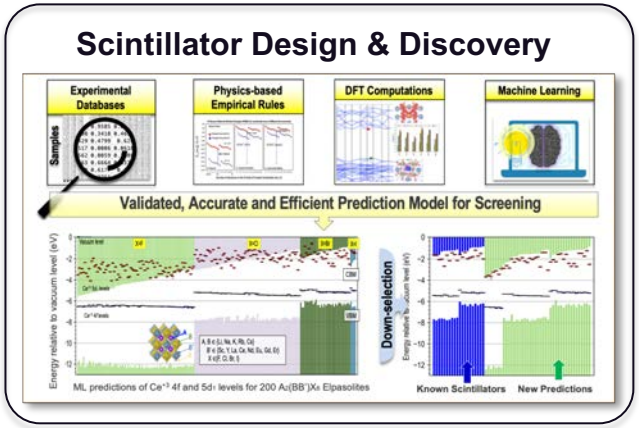


C. Draxl, M. Scheffler, *NOMAD:* ***The FAIR Concept for Big-Data-Driven Materials Science***, MRS Bulletin 43 676 (2018).

Challenges for ML in Materials Science

- **Biased and Heterogeneous Data**
  Chemical intuition driven experiments need not be independent and identically distributed in the target chemical space; results from failed experiments are rarely reported; varying fidelity data sources with different levels of noise/uncertainties
- **Small Datasets**
  High-quality (relevant) data based on experiments and/or quantum mechanical computations can be extremely limited; uncertainty quantification and testing of the trained models becomes critical
- **Domain-knowledge Integration**
  Optimal ways to include domain knowledge of the problem (mechanistic details, fundamental laws, boundary conditions, symmetry relationships etc.) into a ML model may not be obvious
- **Iterative feedback from experiments can be slow**
  Materials design and discovery problems sometime require iterative feedback from experiments in an adaptive design or active learning framework which can be extremely slow bottleneck
- **Conflicting trends among desired properties**
  It is not always possible to simultaneously optimize all desired properties in one material for a given application; Pareto optimal front exploration problems

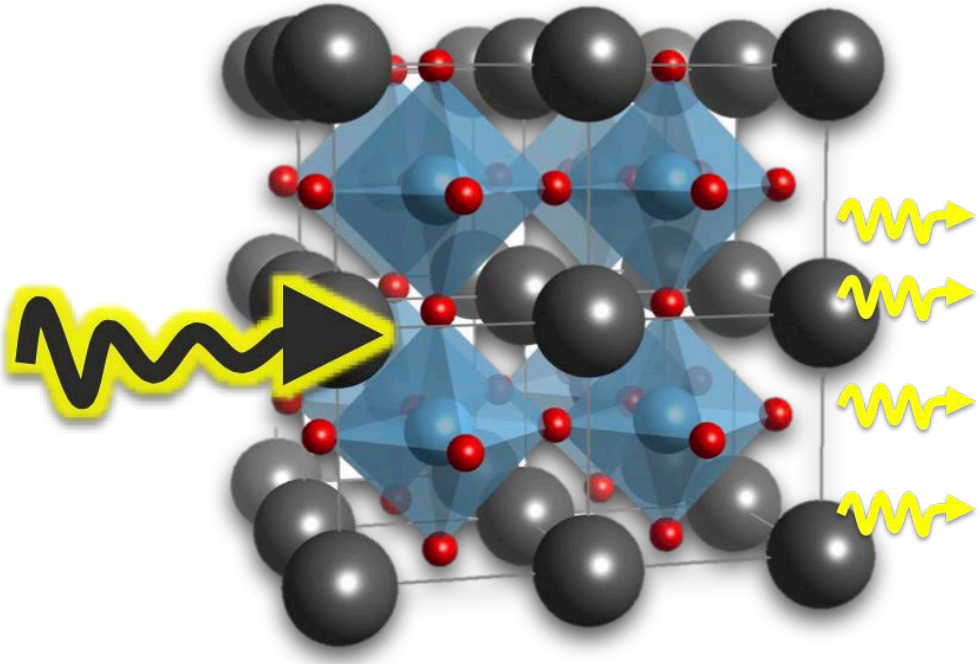
Let’s Get Started…



G. Pilania, S. K. Yadav, M. Nikl, B. P. Uberuaga, and C. R. Stanek, Phys. Rev. Appl. 10, 024026 (2018).
G. Pilania, K. Mcclellan, C. R. Stanek and B. P. Uberuaga, J. Chem. Phys. 148, 241729 (2018).
G. Pilania, X. Y. Liu, and Z. Wang. J. Mater. Sci. 54, 11, 8361 (2019).
A. Talapatra, B. P. Uberuaga, C. R. Stanek, G. Pilania, Submitted to Chem. Mater. (2020).

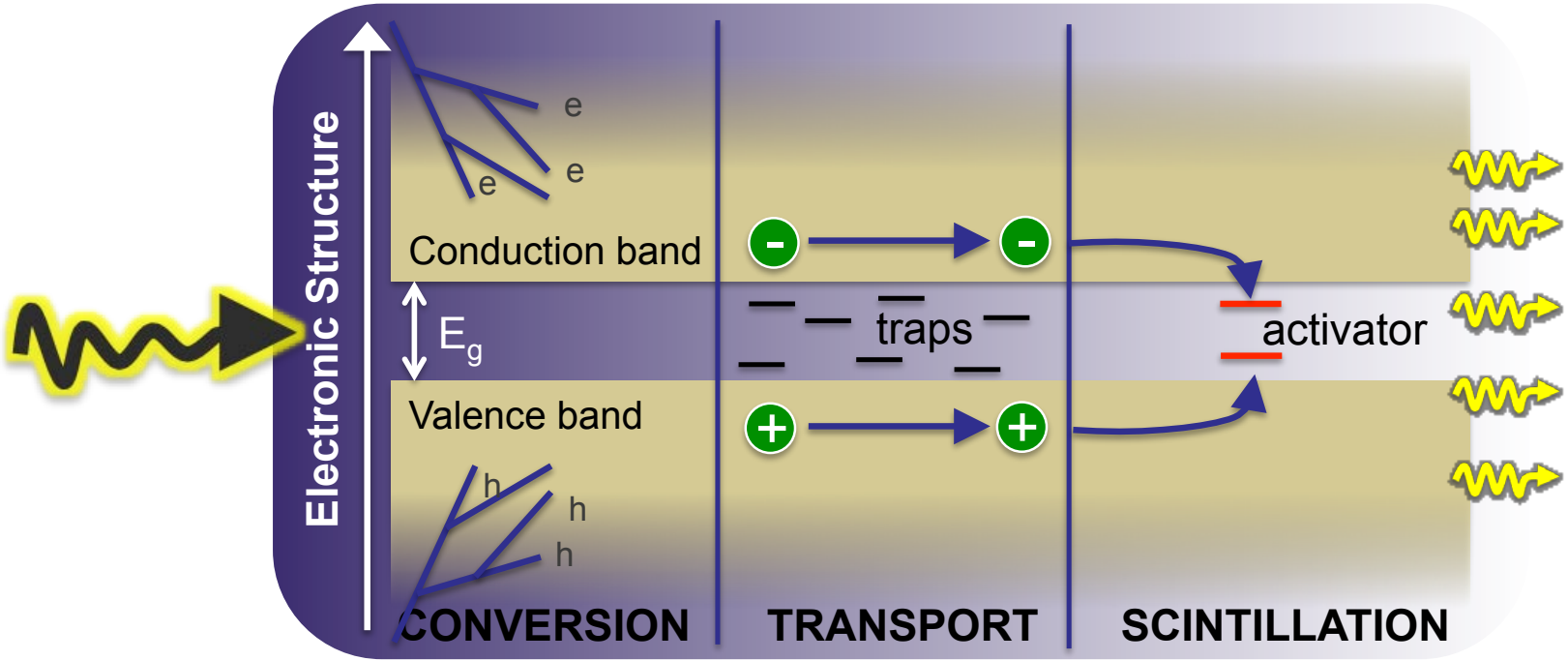
What is a Scintillator?

X- or γ-ray



Visible or
near-visible light

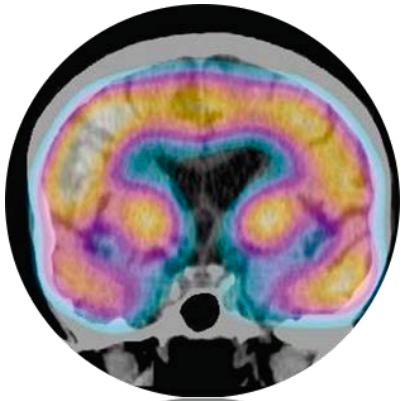
Scintillation is a Complex Process…



A complex multi-step process
Mechanistic details involve a number of materials characteristics

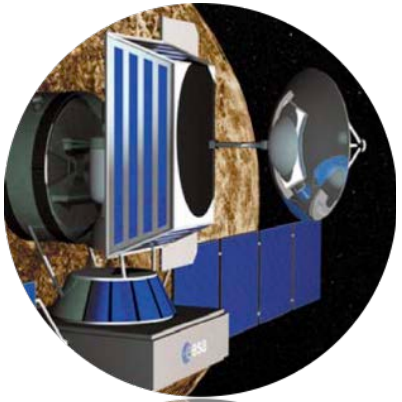
…with a Range of Applications

Global Security & Threat Reduction



Medical Imaging

High Energy Physics Experiments



Space Research

…

ISSUE: in all cases, scintillators are taken "off the shelf"

OPPORTUNITY: Custom Scintillators Design and Development

Desired Scintillator Property Portfolio

High Light Yield
Number of photons per MeV should be high

Fast Response
A single short decay time constant is ideally desired (~ns) with no after glow

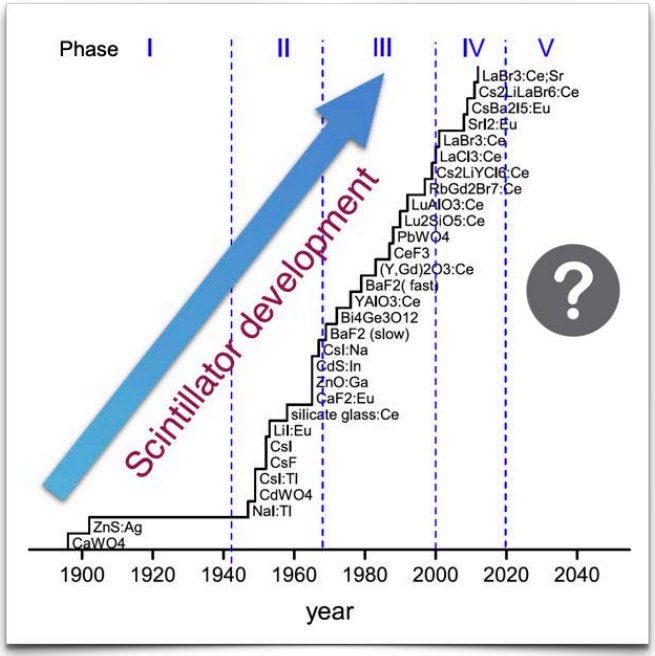
Stability Under High Energy Radiation
In addition to thermodynamic and dynamic stability, the material should be able to withstand high dose of radiation without a significant degradation in performance

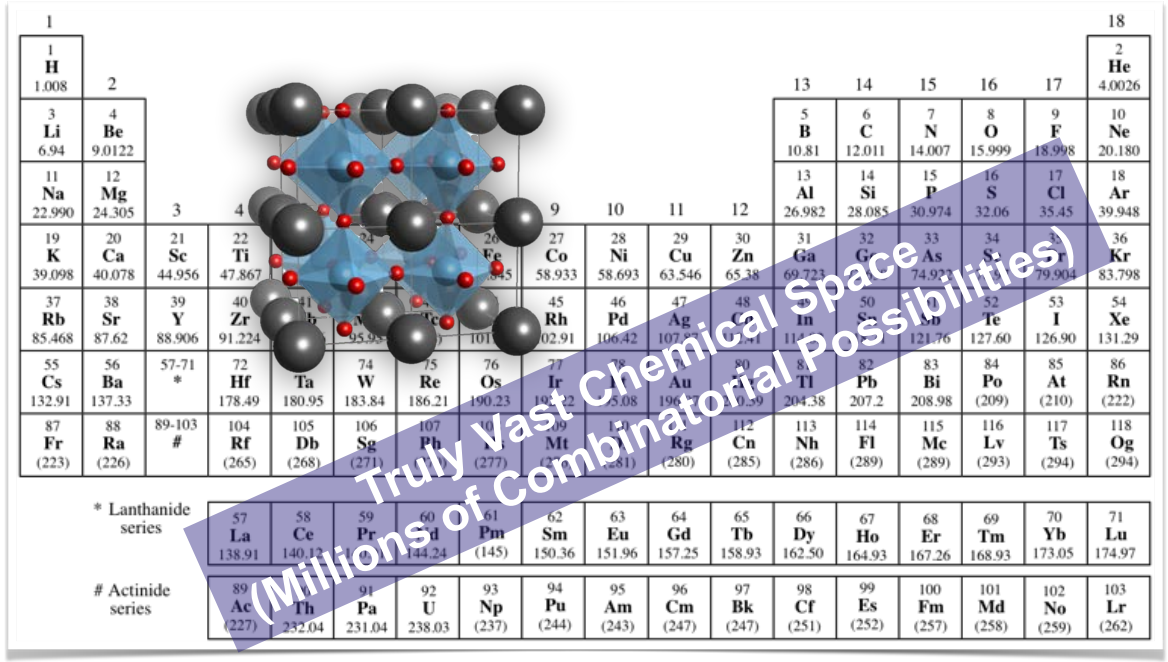
Emission Wavelength
Should be tuned to the specific application and detector system

High Density and Effective Atomic Number
Required for efficient absorption of radiation

Appropriately Placed Activator States; Minimal Deleterious Defects and Trap States
A favorable electronic structure to enable efficient transportation and radiative recombination of charge careers

Discovery of Novel Scintillators

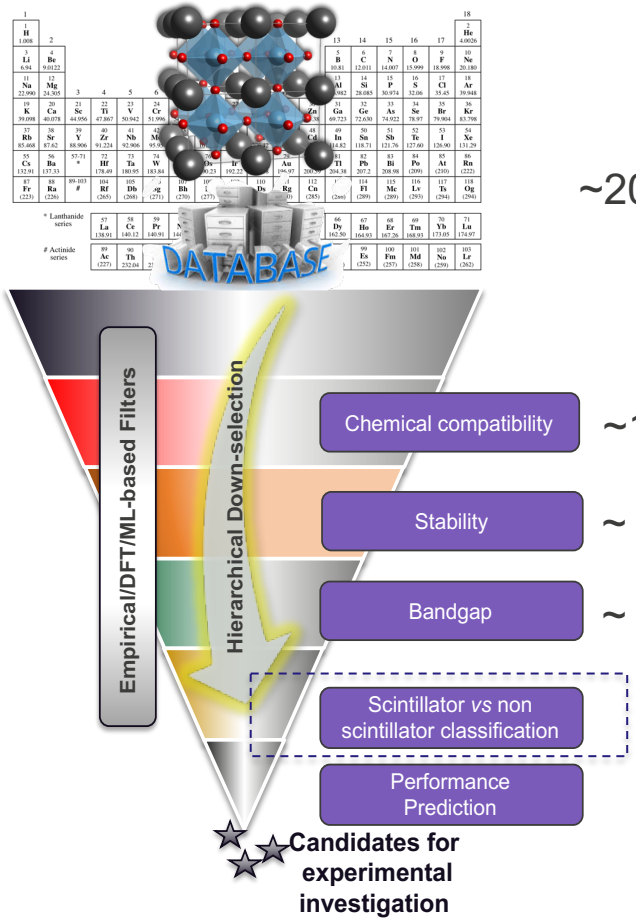




Traditionally, scintillator discovery and development has largely been guided by chemical intuition and laborious trial-and-error based experimentation.
Chemical universe is truly vast and only a tiny fraction has been explored…

P. Dorenbos, The quest for high resolution γ-ray scintillators, Optical Materials: X 1, 100021 (2019)

An Efficient Scintillator Screening Framework



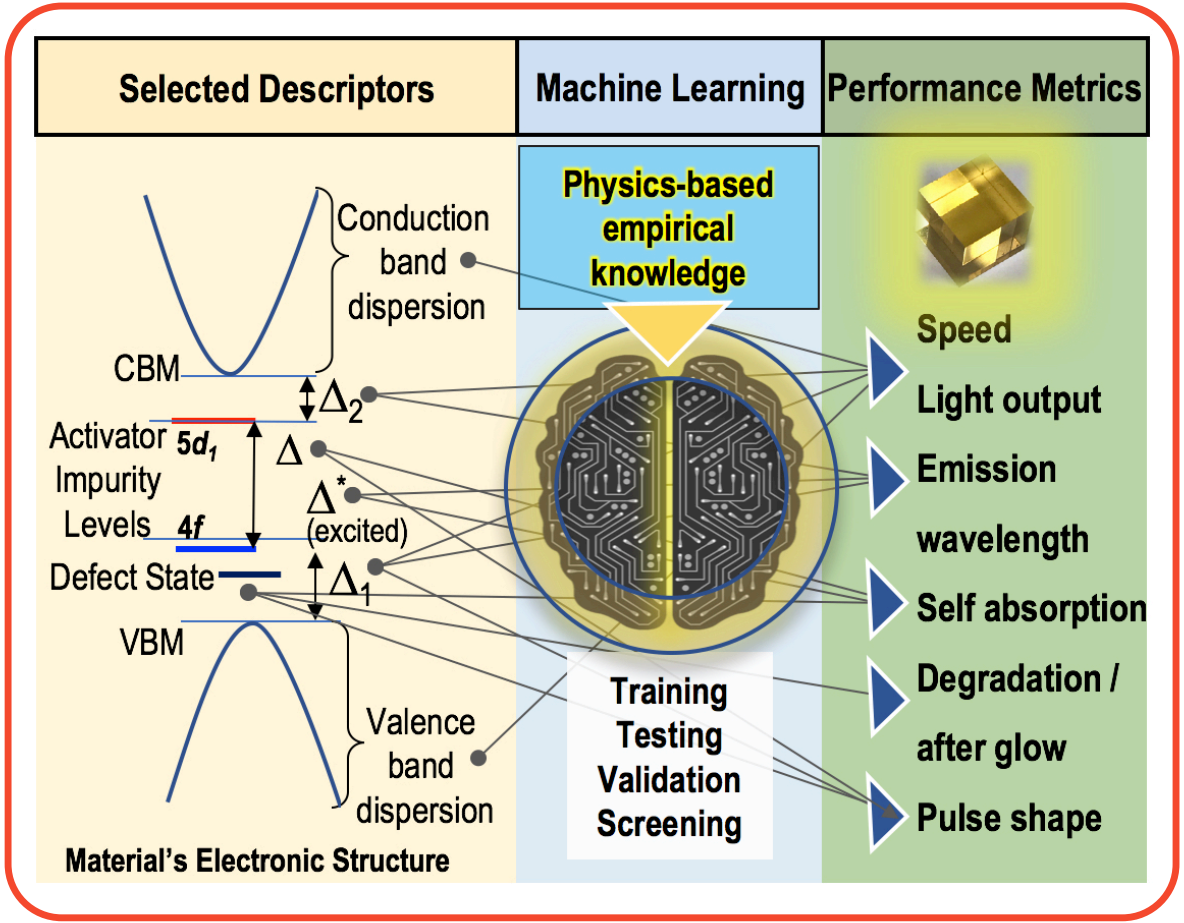
A hierarchical down-selection strategy:

- Starting with a set of ~20M combinatorial possibilities down-select the most promising candidates for experimental investigation
- ML-based models for perovskite formability, thermodynamic stability, bandgap and scintillation performance
- Once potential scintillator chemistries have been identified, ML-based models can be used to predict performance metrics

Machine Learning Enabled Mappings

"Models of Performance"

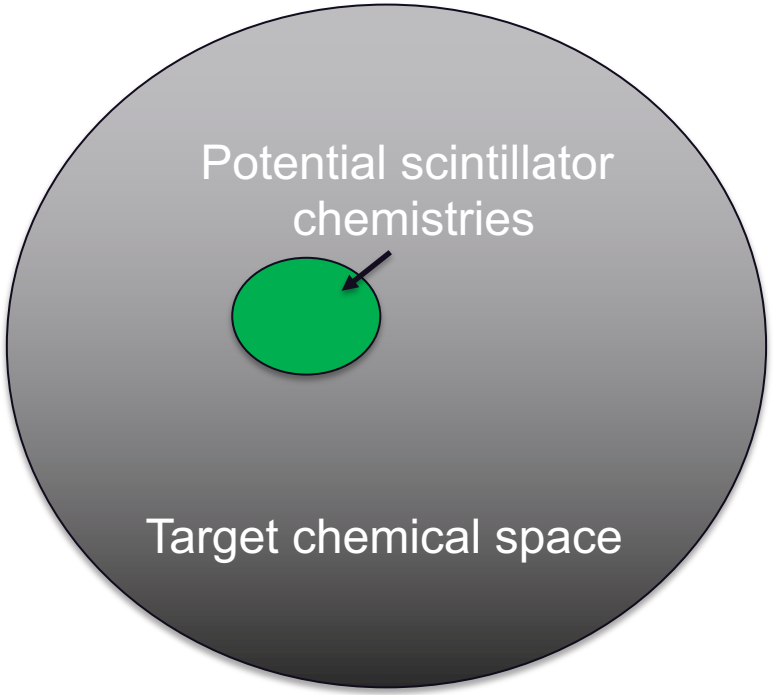
- **A naive traditional approach fails completely**
- **An ML model trained only on scintillators would be inadequate for predictions**
  - Most of the compounds tend to be non-scintillating
  - The trained model would still predict a finite light output, for instance



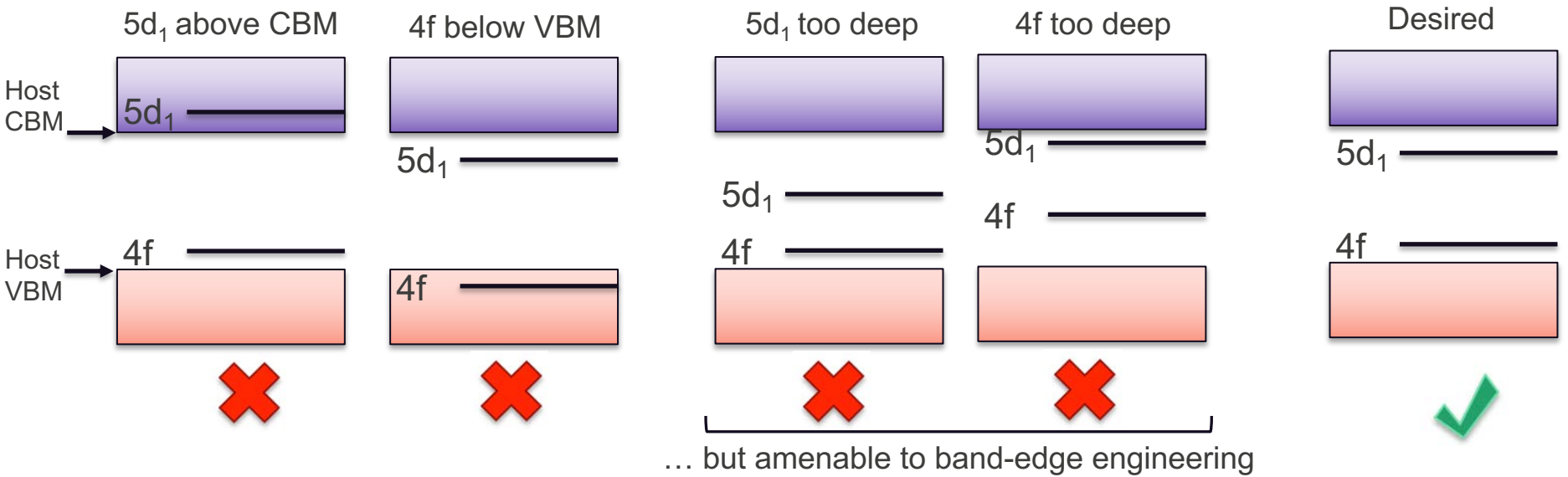
Machine Learning Enabled Mappings

Early Lessons Learned:

- ML-based “models of performance” can not be used directly to explore large chemical spaces
- A scintillator versus non-scintillator classification problem needs to be addressed first
- Once potential scintillator chemistries have been identified, ML-based models can be used to predict performance metrics



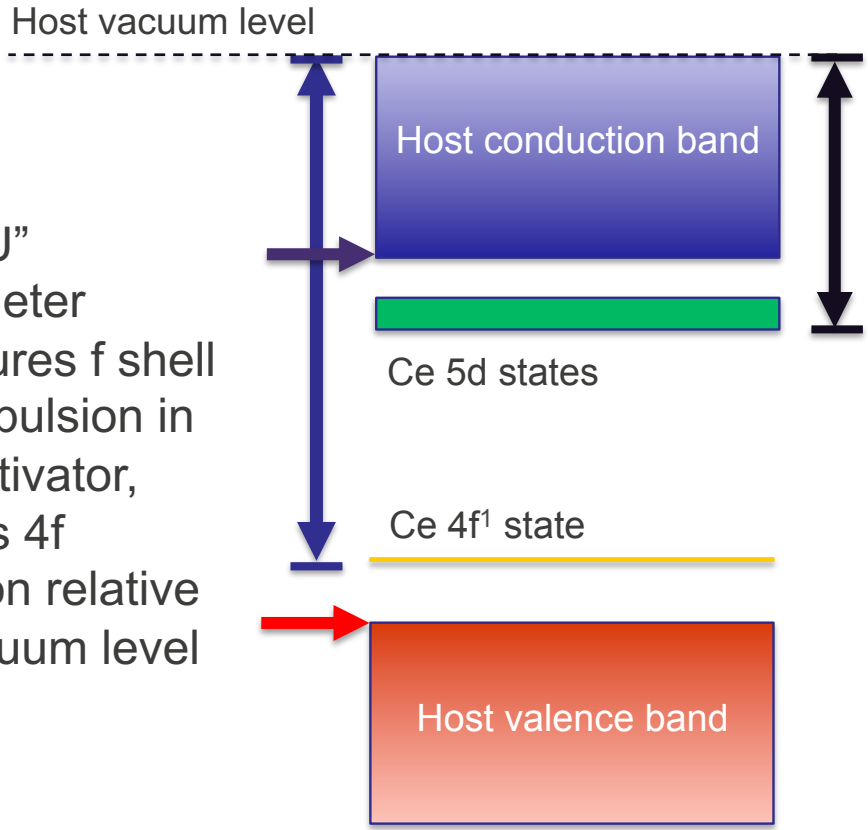
Identification of Potential Scintillators



In Ln-doped inorganic scintillator chemistries, position of the activator levels with respect to the band edges can be used for a high throughput screening.

What governs location of lowest *4f* and *5d* levels with respect to the band edges?

Scintillator vs Non-scintillator: The “U” and “D” parameters



The “U” parameter measures f shell e-e repulsion in the activator, relates 4f position relative to vacuum level

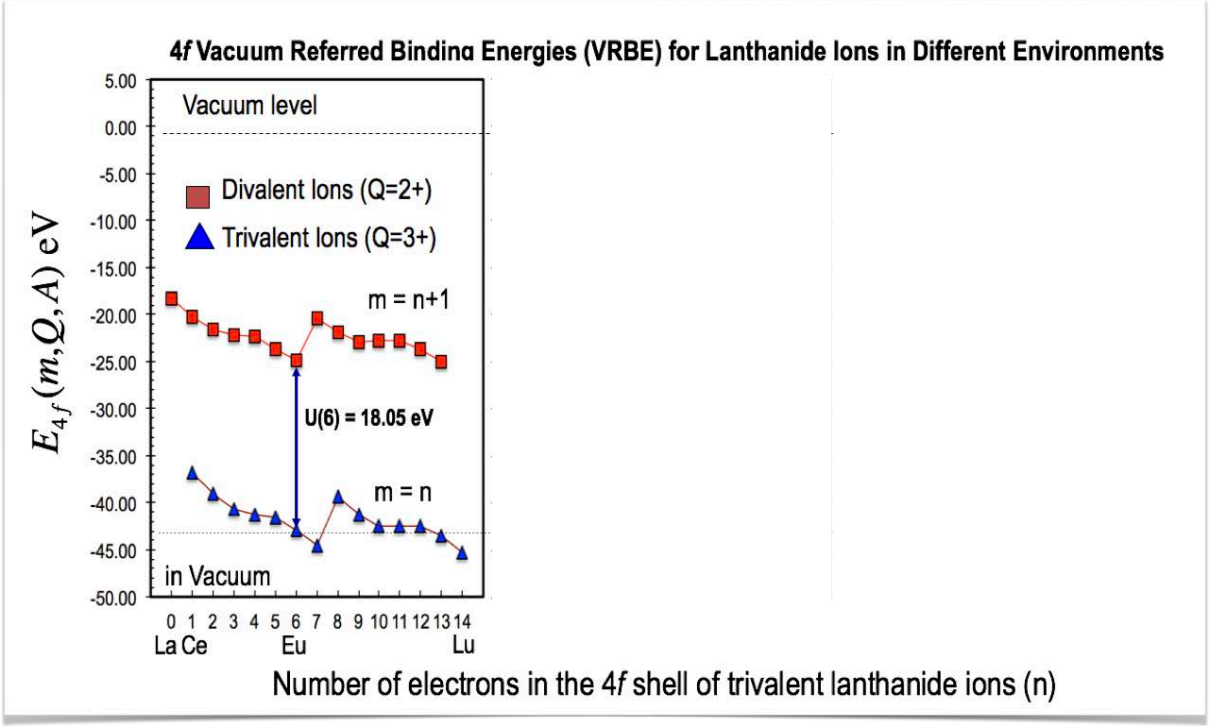
The “D” parameter is primarily governed by the crystal field around the activator

DFT-based electronic structure computations for the vacuum referenced band edges of the host chemistry

‡P. Dorenbos, Physical Review B 85, 165107 (2012).

A closer Look at the “U” Parameter

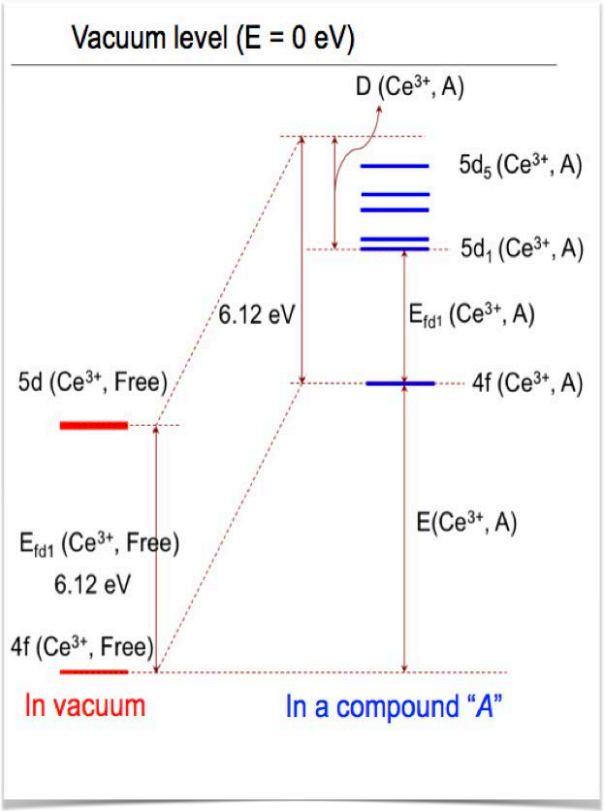
What governs location of 4*f* energy levels in Ln ions?



P. Dorenbos, Physical Review B 85, 165107 (2012).

The “U” parameter is a measure of f shell e-e repulsion in the lanthanide activator

A closer Look at the “D” Parameter



The “D” parameter is primarily governed by host’s dielectric screening and crystal field around the dopant/activator site

A knowledge of the “**U**” and “**D**” parameters combined with **Dorenbos’ chemical shift model** and **DFT**-based electronic structure computations for the bandgap allows for an accurate prediction of the 4f and $5d_1$ levels of a lanthanide activator in a given host chemistry.

P. Dorenbos, Physical Review B 85, 165107 (2012).

Details of the Training Dataset for ML

Training Data
for the “U” & “D” Parameters

NH

ELSEVIER

Journal of Luminescence 91 (2000) 155–176

JOURNAL OF LUMINESCENCE

www.elsevier.com/locate/jlumin

The 5d level positions of the trivalent lanthanides in inorganic compounds

P. Dorenbos*

Radiation Technology Group, Interfaculty Reactor Institute, Delft University of Technology, Mekelweg 15, 2629 JB Delft, The Netherlands

Journal of Luminescence 135 (2013) 93–104

Contents lists available at SciVerse ScienceDirect

Journal of Luminescence

journal homepage: www.elsevier.com/locate/jlumin

ELSEVIER

LUMINESCENCE

Ce^{3+} 5d-centroid shift and vacuum referred 4f-electron binding energies of all lanthanide impurities in 150 different compounds

Pieter Dorenbos*

Luminescence Materials Research group, Department of RRR, Faculty of Applied Sciences, Delft University of Technology, Mekelweg 15, 2629 JB Delft, Netherlands

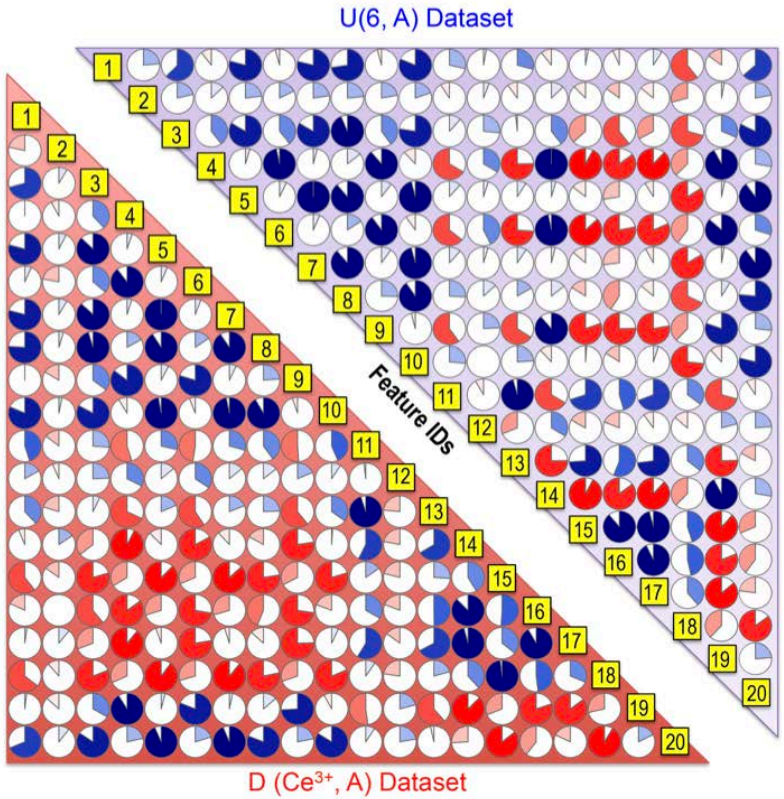
Feature set
Elemental & bulk features of host chemistries

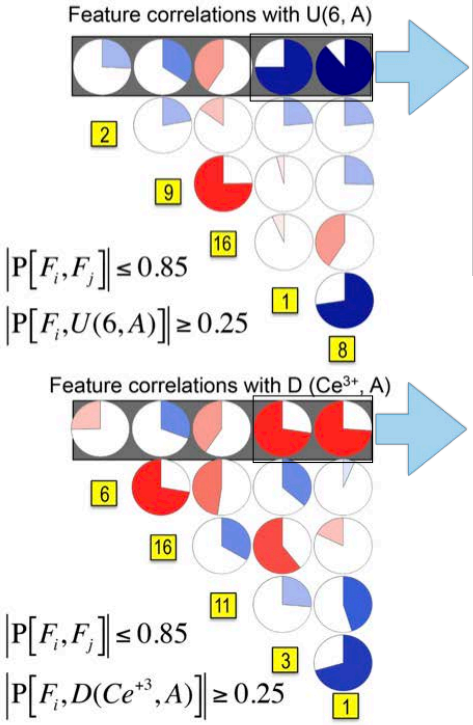
Feature ID	Symbol	Description	Feature ID	Symbol	Description
1	E_g	Bandgap (computed within DFT using the PBE functional)	11	EA_{av}	Average electron affinity
2	ρ	Density of host compound	12	$EA(+)_{av}$	Average electron affinity for cationic species
3	χ_{av}	Average electronegativity	13	$EA(-)_{av}$	Average electron affinity for anionic species
4	$\chi(+)_{av}$	Average electronegativity for cationic species	14	$r(+)_{av}$	Average empirical radius for cationic species
5	$\chi(-)_{av}$	Average electronegativity for anionic species	15	$r(-)_{av}$	Average empirical radius for anionic species
6	$\chi(+)^q_{av}$	Charge weighted average electronegativity for cationic species	16	P_{av}	Average atomic polarization
7	$\chi(-)^q_{av}$	Charge weighted average electronegativity for anionic species	17	$P(+)_{av}$	Average atomic polarization for cationic species
8	IP_{av}	Average ionization potential	18	$P(-)_{av}$	Average atomic polarization for anionic species
9	$IP(+)_{av}$	Average ionization potential for cationic species	19	$M(+)_{av}$	Average Pettifor’s Mendeleev’s number for cationic species
10	$IP(-)_{av}$	Average ionization potential for anionic species	20	$M(-)_{av}$	Average Pettifor’s Mendeleev’s number for anionic species

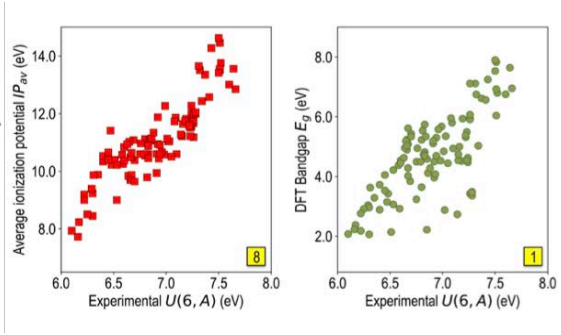
P. Dorenbos, Journal of Luminescence 135, 93 (2013).
P. Dorenbos, Journal of Luminescence 91, 155 (2000).

Except Bandgap and density of the host, all other features are based on the properties of atomic constituents.

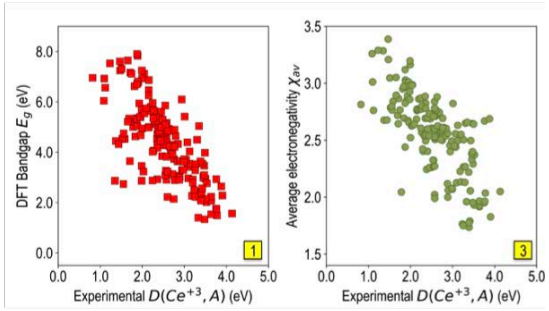
Pairwise Feature Correlations







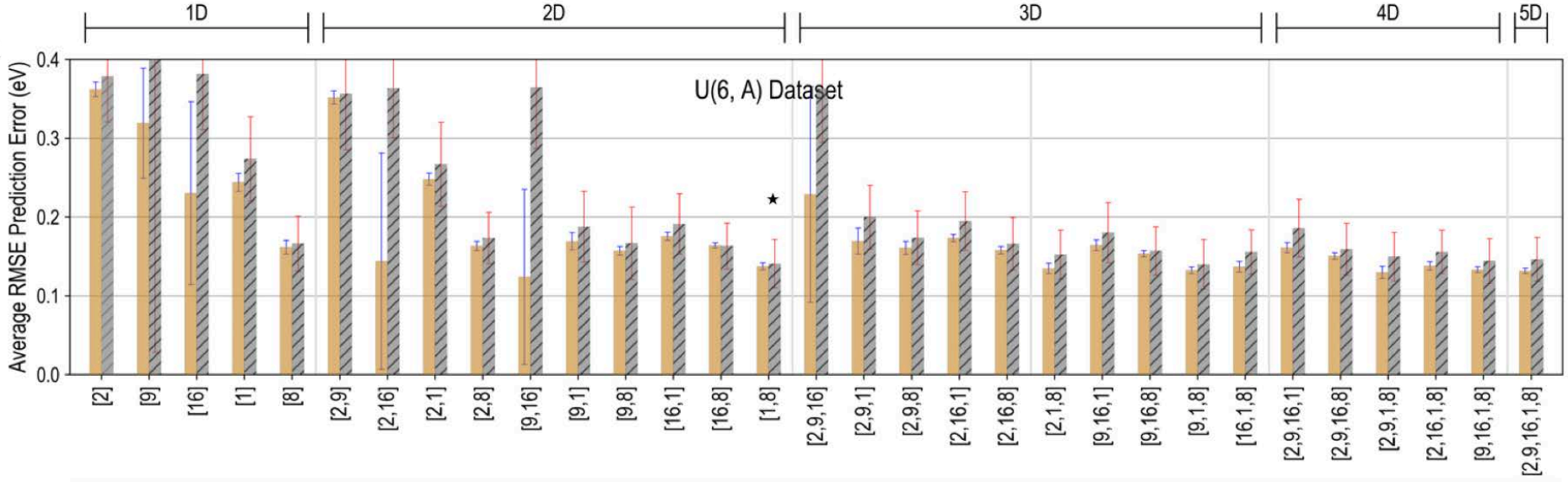
Top Features for the "U" Model



Top Features for the "D" Model

G. Pilania, et al. J. Chem. Phys. ***148*** *241729 (2018).*

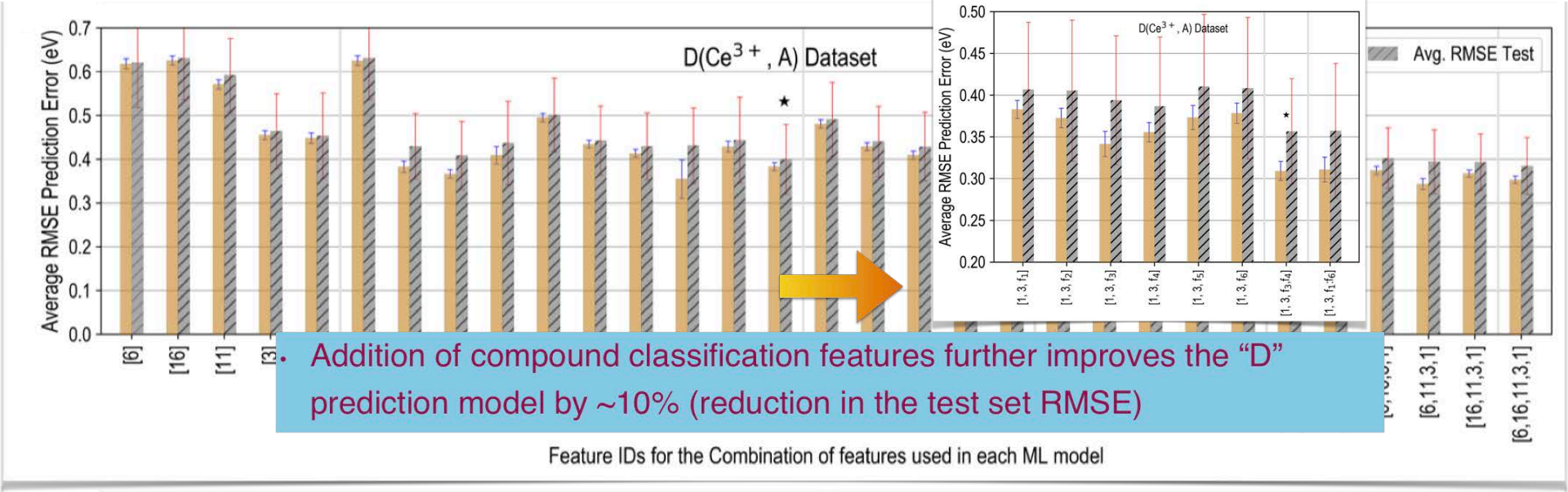
The “U” Parameter: ML Prediction Performance



1: PBE Bandgap [E_g], 2: Host density [$\boldsymbol{\rho}$], **3**: Avg. Electronegativity [$\boldsymbol{\chi}_{av}$], **6**: Charge weighted avg. electronegativity of cations [$\boldsymbol{\chi}(+)^{q}_{av}$], **8**: Avg. ionization potential [IP_{av}], **9**: Avg. ionization potential for cations [$IP(+)_{av}$], **11**: Avg. electron affinity [EA_{av}], **16**: Avg. atomic polarization [P_{av}]

Can predict the “U” parameter to within 0.1-0.2 eV
Best features are PBE bandgap and average ionization potential

The “D” Parameter: ML Prediction Performance



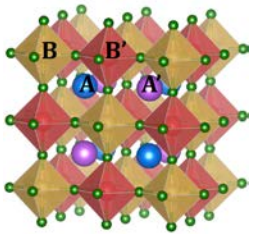
- Addition of compound classification features further improves the “D” prediction model by ~10% (reduction in the test set RMSE)

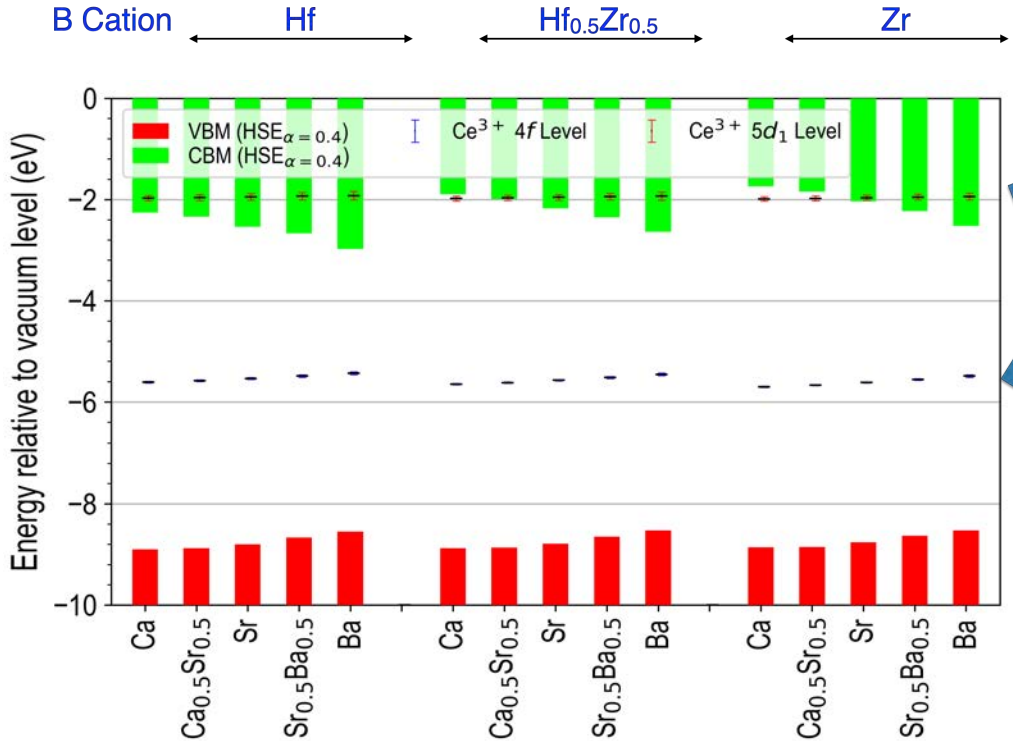
Feature IDs for the Combination of features used in each ML model

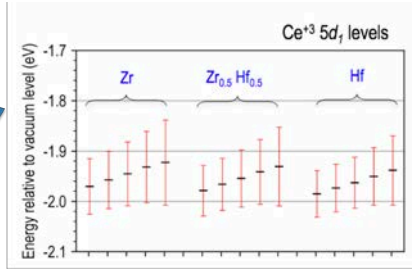
1: PBE Bandgap [E_g], 2: Host density [ρ], **3**: Avg. Electronegativity [χ_{av}], **6**: Charge weighted avg. electronegativity of cations [$\chi(+)^q_{av}$], **8**: Avg. ionization potential [IP_{av}], **9**: Avg. ionization potential for cations [$IP(+)_{av}$], **11**: Avg. electron affinity [EA_{av}], **16**: Avg. atomic polarization [P_{av}]

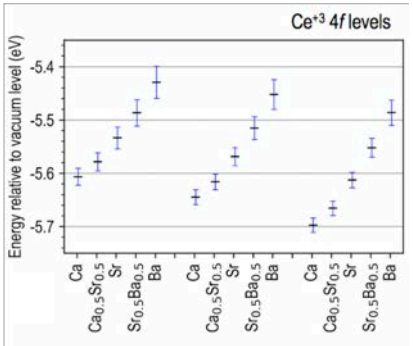
Can predict the “D” parameter to within ~0.3 eV
Best features are PBE bandgap and average electronegativity

Chemical Trends are Well Captured







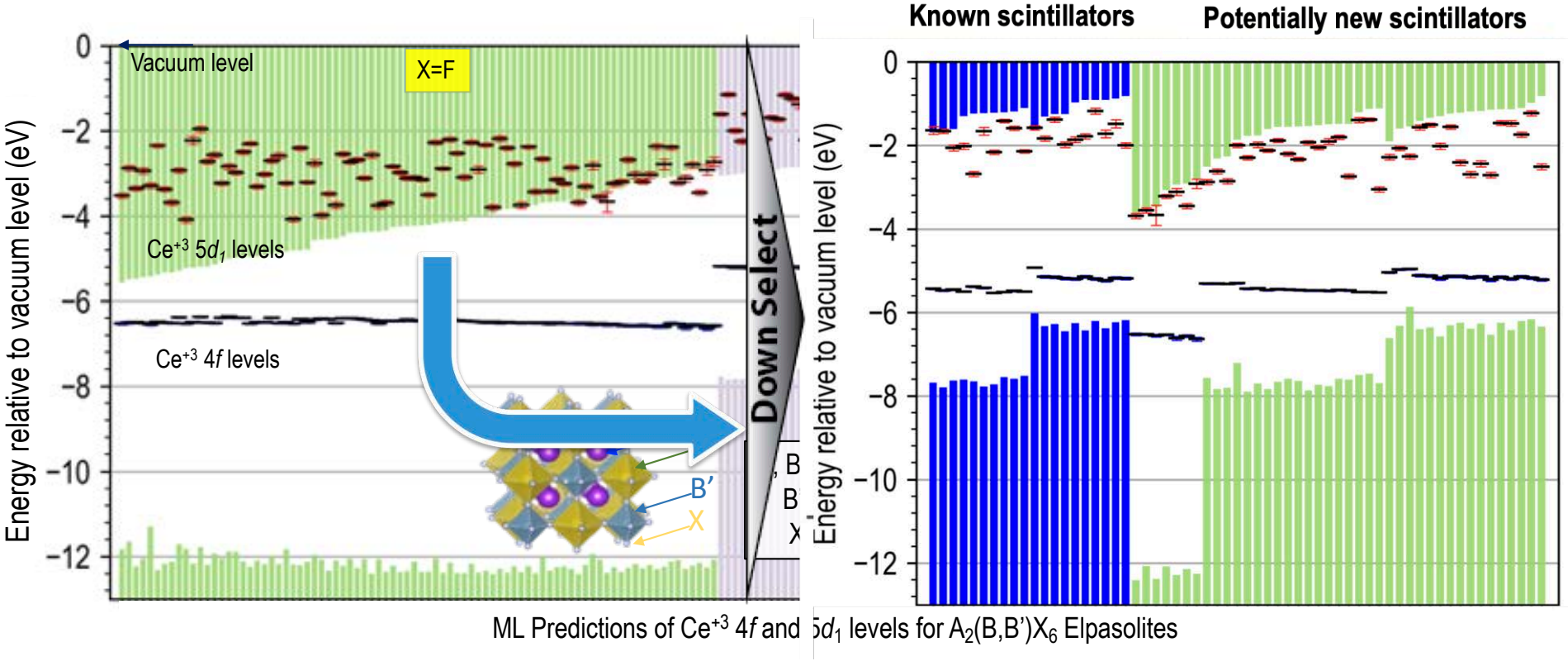


G. Pilania, et al. J. Chem. Phys. ***148*** *241729 (2018).*

P. Dorenbos, Journal of Luminescence 151, 224 (2014).

How Predictive this Approach is?

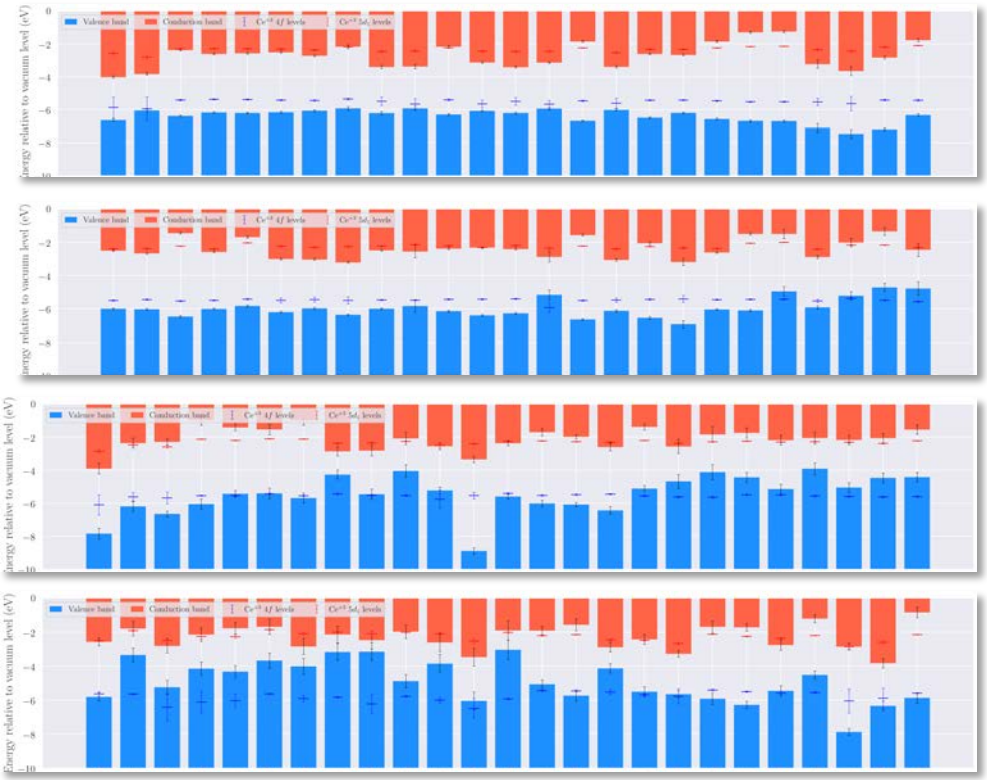
Application to Elpasolites



G. Pilania, et al. J. Chem. Phys. ***148*** *241729 (2018).*

Ready to Make Predictions on Double Perovskites!

Example predictions on 100 double perovskite oxides



Additional Considerations:

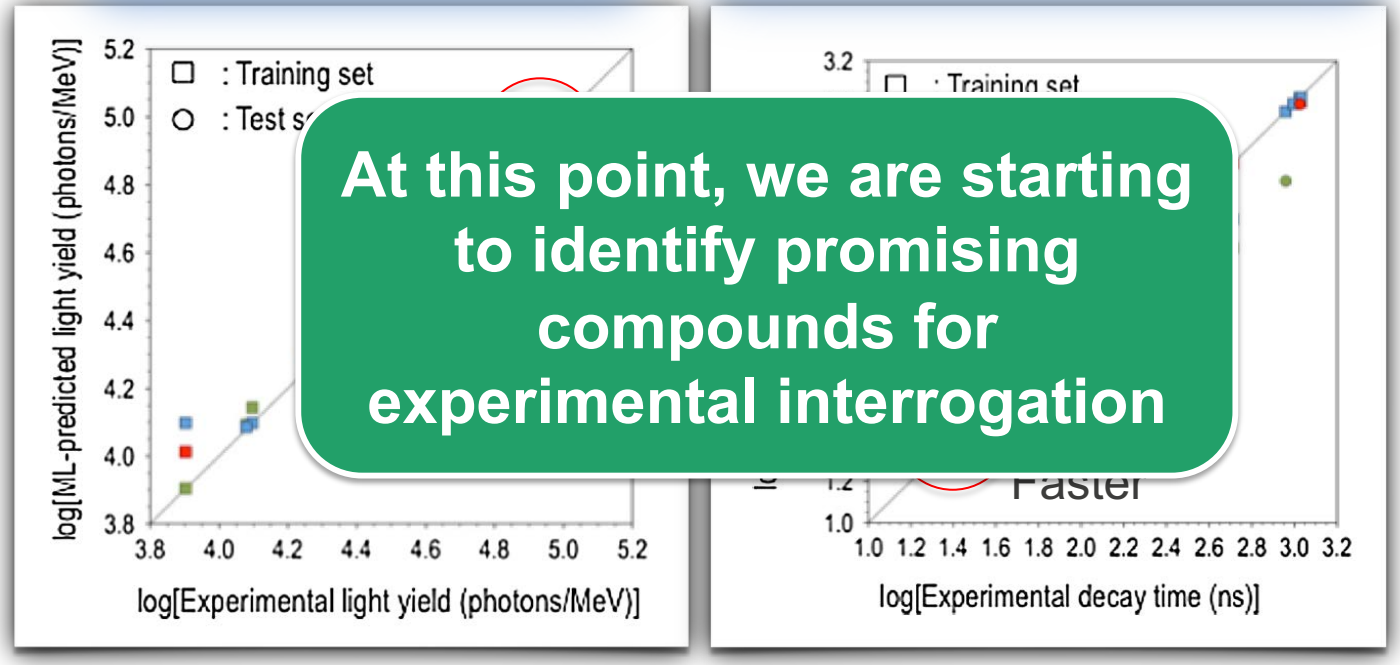
- Compounds with larger tolerance factors for isotropic optical properties
- Prefer compounds with non-lanthanide ions to avoid non-radiative recombination pathways triggered by the *4f*-*4f* transition manifolds
- Remove compounds with toxic volatile ions and screen for improved synthesizability
- Prefer compounds with higher density and effective atomic number
- Prefer +3/+5 B-site double perovskites due to improved substitutional dopability of Ce^{+3} activator
- Account for cost and availability factors

Models of Performance for Scintillators

Looking only for qualitative predictions at this point

Light Yield

Fastest Decay time Const.



At this point, we are starting to identify promising compounds for experimental interrogation

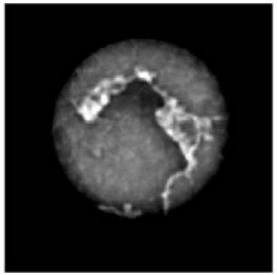
Top-Down Approach: Building Models from Experimental Data

- A range of Ce and Eu doped scintillators from experiment
- Not looking for quantitative accuracy → qualitative predictions are good enough for screening

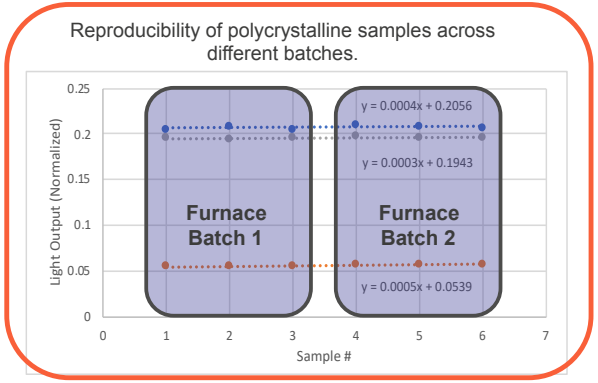
G. Pilania, et al. J. Mater Sci. ***54*** *8361 (2019).*

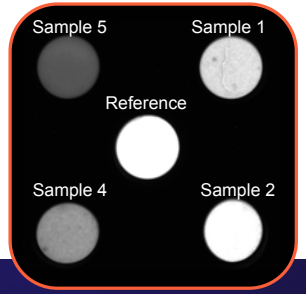
Challenges with Ceramics Samples

- Turnaround time with single crystal synthesis and measurements is too long.
- Resorting to ceramic pellets
  - with no sintering: poor mechanical strength, porosity, voids
  - with strong sintering possibility of Ce segregation to grain boundaries
  - Reproducibility issues
- Optimized processing conditions with lightly sintered pellets lead to reproducible relative light yield measurements



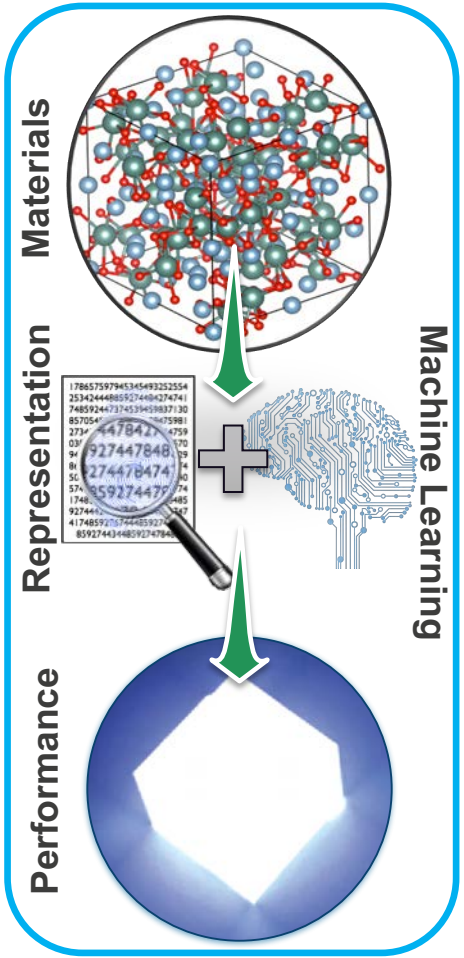
Poor mechanical strength, cracks, and porosity can make relative light yield measurements difficult





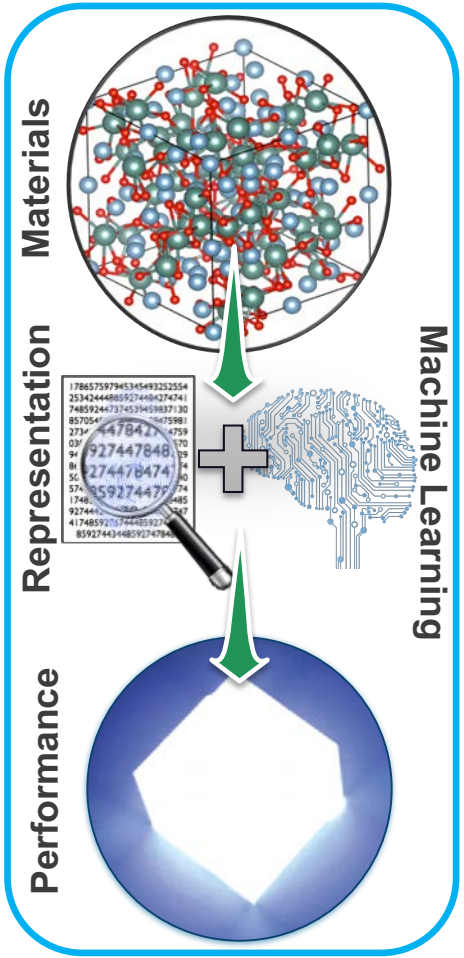
Reflecting Back on the Challenges

- **Biased and Heterogeneous Data**
  Highly exclusive and expert-curated datasets were used; data from single crystals and ceramics samples is not mixed
- **Small Datasets**
  Uncertainty quantification via boot strapping; simplest possible models were favored for better generalizability; causal relationships of the selected features with the target property was rationalized
- **Domain-knowledge Integration**
  Two regression models were developed to “solve” a classification problem—guided by the underlying physics of the problem
- **Iterative feedback from experiments can be slow**
  Resorting to polycrystalline ceramic samples for the screening stage of the project for a faster experimental feedback
- **Conflicting trends among desired properties**
  Light output and speed are inversely correlated; looking for application-specific optimal candidates rather than “the best” scintillator chemistry



Conclusions

- Scintillator screening was used as an example to demonstrate that domain-knowledge-informed machine learning can expedite materials design and discovery
- Challenges unique to such a data-enabled approach were highlighted and discussed
- Physics-informed scintillator versus non-scintillator classification problem was solved
- The classification model was validated via predicting chemical trends and predicting known Elpasolites scintillators
- The developed ML-based hierarchical down-selection approach has implications beyond scintillation applications



Acknowledgements

The Team


Ghanshyam Pilania


Anjana Talapatra


Chris Stanek


Ben Liu


Blas Uberuaga

Modeling


Ken McClellan


Jan Bárta


Kim Pestovich

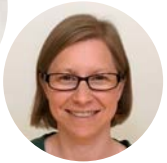

Mandie Gehring


Caleb Chandler

Making


Todd Haines


Brenden Wiggins


Jessica Clayton (NNSS)


Rod McCabe


James Valdez


Erick Smith

Measuring

LANL Computational Resources & Funding