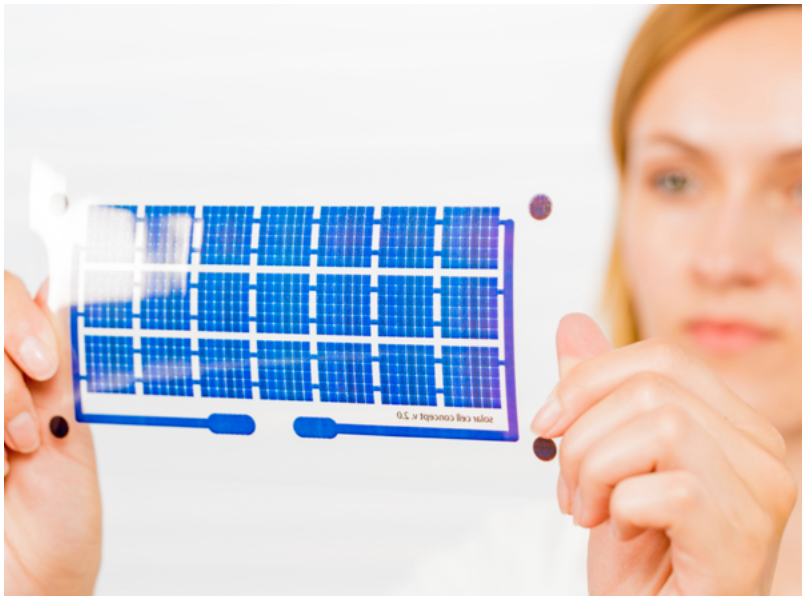
CITRINE INFORMATICS
citrine.io • info@citrine.io

MACHINE LEARNING ACCELERATES MATERIALS DISCOVERY

PANASONIC DEVELOPS NEW, PATENT PENDING ORGANIC SEMICONDUCTOR MOLECULES WITH CITRINE'S AI PLATFORM

CITRINE INFORMATICS Panasonic



EXECUTIVE SUMMARY

IP GROWN
PATENTS PENDING ON HIGH MOBILITY THIEONOACENES

PERFORMANCE IMPROVED
25% INCREASE IN HOLE MOBILITY CALCULATED

TIME SAVED
ONLY 196 DFT CALCULATIONS PERFORMED ON > 1 MILLION CANDIDATES

NEW INSIGHTS
INTO HOW MOLECULAR TOPOLOGY AFFECTS HOLE MOBILITY

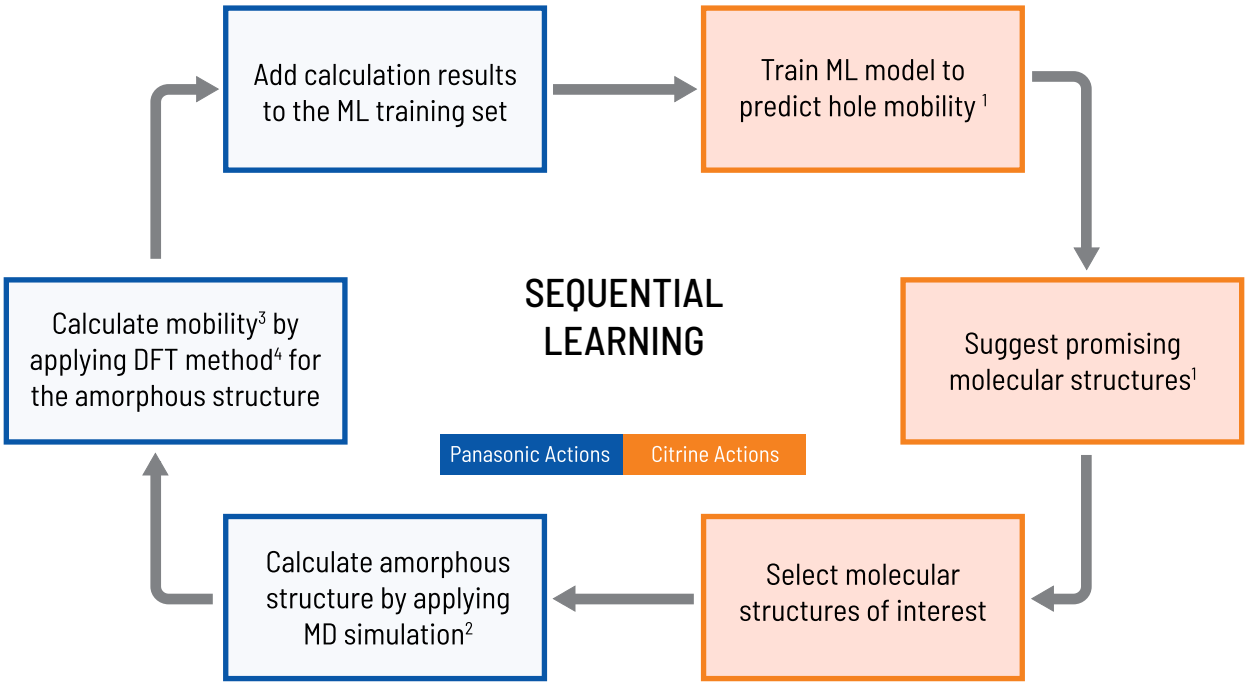
> *"This work demonstrates the utility of using the sequential learning methodology to design experiments for the discovery of novel materials."*
>
> Nobuyuki Matsuzawa,
> **Panasonic**

THE CHALLENGE

The internet of things requires flexible, lightweight and inexpensive semiconductors. Panasonic is already using organic semiconductors for IOT applications, but they needed a way to reduce processing costs while also increasing performance. Heteroacenes are a molecular class of organic semiconductor that are soluble in organic solvents, enabling cheap spin-coating techniques. But there are millions of them, far too many to exhaustively survey experimentally or computationally. Panasonic and Citrine combined their expertise to meet this challenge.

THE APPROACH

Although computing power is increasing every day, Density Functional Theory (DFT) and Molecular Dynamics (MD) simulations can still take days to perform. Panasonic decided that using Machine Learning (ML) to spot promising molecular structures would save time and focus their simulation efforts on candidates with a high likelihood of success. By using Machine Learning, the Panasonic team evaluated the effect of molecular structure on material performance on many more candidates than would've been possible using physics-based simulation methods. This increased the chance of finding high-performing candidates and delivering more value to the business.



ITERATIVE MOLECULAR DESIGN WITH SEQUENTIAL LEARNING

The initial machine learning model was trained on only 32 DFT calculated results. The model used molecular structure – represented by a Simplified Molecular Input Line Entry System (SMILES)[5,6] string – as an input to predict electron and hole mobility.The team expected the intial model to have low accuracy and high uncertainty due to a small training dataset. They then leveraged this uncertainty to identify which DFT calculations to perform next, performed the simulations, and updated the training set with the new data. This iterative process, called sequential learning, allowed the team to increase the predictive quality of the model over time.

Two types of experiments were identified: A) candidates where the model is highly uncertain (Maximum Uncertainty (MU))– so that it can strengthen the model in weak areas and B) candidates where there is the highest expected hole mobility (Maximum Expected Improvement (MEI)), so that initial predictions of high-performing materials can be evaluated.

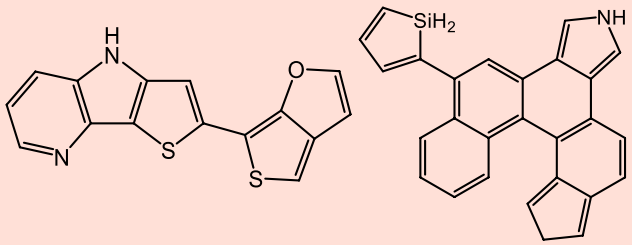
citrine.io • info@citrine.io

DESIGN SPACE

The Design Space refers to the set of possible materials to be evaluated. For this project, we represented the design space with a finite set of molecular structures. Previous research found high hole mobilities in fused thiophenes and selenophenes, and explored the effect of topological shape on hole mobility. However, it has been difficult to uncouple the nature of single molecules and intermolecular effects. Initially a subset of the Harvard Clean Energy Project database of candidate solar materials was used as the design space.

INITIAL DESIGN SPACE

Subset of Harvard Clean Energy Project database of candidate organic solar materials[1] (~130,000 candidate structures)

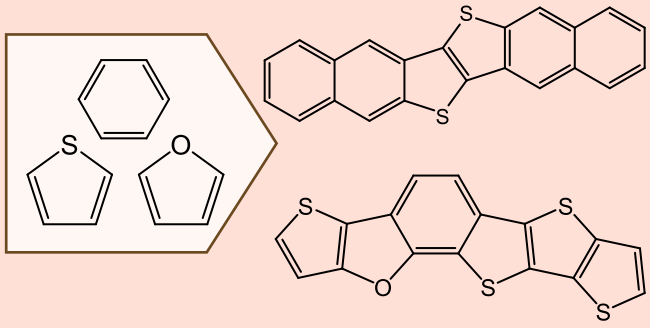


FEATURE ENGINEERING

Citrine produced custom features, also called descriptors, to more accurately represent the fused ring structures of the target molecules. The new features represent information such as the curvature of the molecule and the number of rings present. Features like this are developed by Citrine to increase ML model performance by incorporating materials science and chemistry domain knowledge. Citrine's platform enabled the Panasonic scientists to see which features had the highest correlation with their target properties, which helped them learn from initial iterations and create a custom design space. The team chose a set of molecular building blocks (benzene, furan, and thiophene) and a set of rules about how they could be combined. Citrine's materials-aware AI platform quickly screened 250 million possible molecules down to a custom design space consisting of 500,000 viable candidates.

CUSTOM DESIGN SPACE

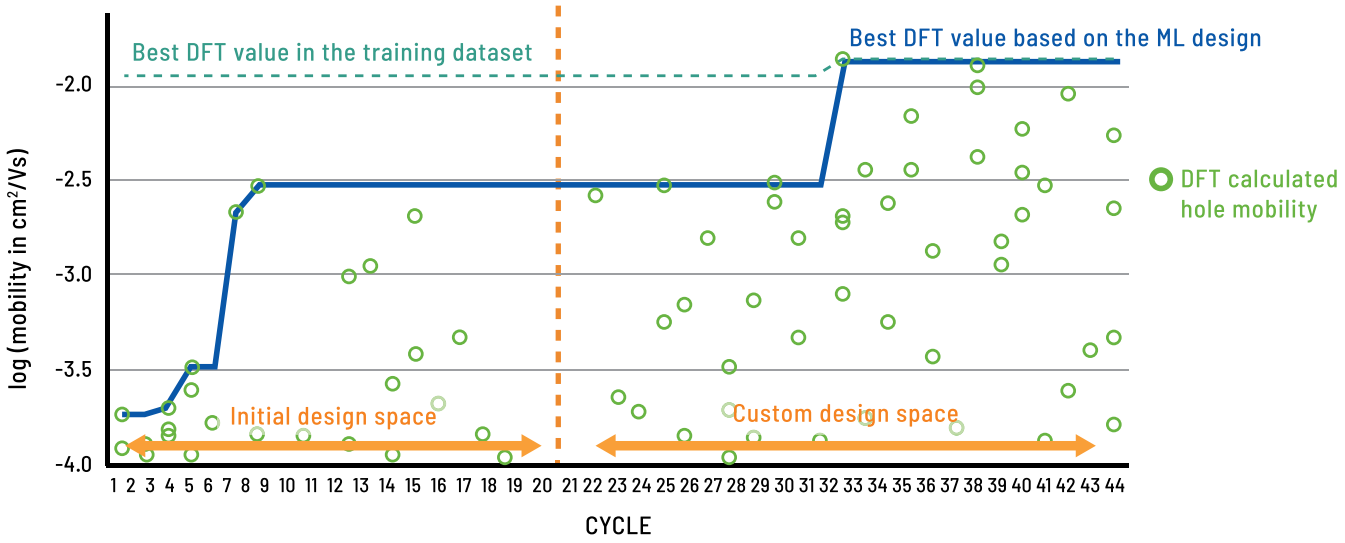
Structures based on fused thiophenes and furans, constructed from aromatic-ring building blocks (~500,000 candidate structures)

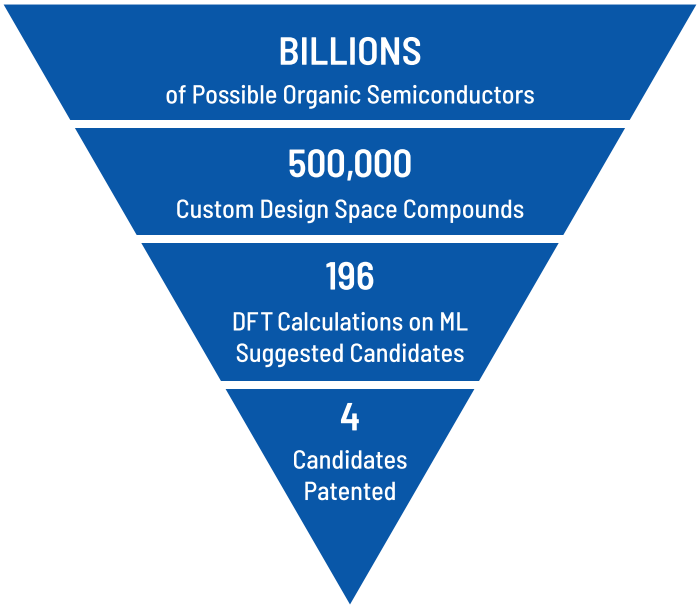


citrine.io • info@citrine.io

THE RESULTS

This AI-driven research methodology helped the Panasonic team discover 4 new molecules, one of which had a 25% higher hole mobility than previously known molecules. In addition, it helped Panasonic discover a new scientific insight; compounds with more aromatic rings – half of them being thiophene rings – have higher hole mobility. Four of these compounds are patent pending and planned for synthesis. This sequential learning methodology allowed the Panasonic team to focus their attention on high potential candidates and speed up their R&D efforts. They also learned how molecular topology influences hole mobility at a fraction of the time and cost of exhaustively evaluating the design space.





REFERENCES

1: Ling, J., Hutchinson, M., Antono, E., Paradiso, S., & Meredig, B. (2017). High-Dimensional Materials and Process Optimization using Data-driven Experimental Design with Well-Calibrated Uncertainty Estimates. Integr Mater Manuf Innov 6: 207.
2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2016. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2016.
3: Evans D R, Kwak H S, Giesen D J, Goldberg A, Halls M D, Oh-e M, 2016 Estimation of charge carrier mobility in amorphous organic materials using percolation corrected random-walk model, Org. Electronics 29, 50.
4: Bochevarov A D, Harder E, Hughes T F, Greenwood J, Braden D, Rinaldo D, Halls M D, Friesner R A 2013 Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences, Int. J. Quantum Chem. 113 2110. The software is available from Schrödinger, New York.
5. Weininger, D., Weininger, A. & Weininger, J. L. SMILES 2. Algorithm for generation of unique SMILES notation. J. Chem. Info. Comput. Sci. 29, 97-101 (1989).
6. Citrine Public Features Descriptions. https://help.citrination.com/knowledgebase/articles/1863853- citrine-public-feature-descriptions.
7. Hachmann, Johannes, et al (2011). The Harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid. J. Phys. Chem. Lett. : 2241-2251.

FIND OUT MORE

A paper in the Journal Chemistry of Materials can give you more technical detail on this project.

ARRANGE AN ONLINE MEETING

DOWNLOAD MORE CASE STUDIES

LISTEN TO OUR PODCASTS

SUBSCRIBE TO OUR NEWSLETTER

ABOUT CITRINE INFORMATICS

Citrine Informatics is the award-winning materials informatics platform for data-driven materials and chemicals development. It won the 2017 World Materials Forum Start-up Challenge, the 2018 AI Breakthrough award as the "Best AI-based Solution for Manufacturing", and 2020 Cleantech 100 honors. The Citrine Platform combines smart materials data infrastructure and Artificial Intelligence, which accelerates development of cutting-edge materials, facilitates product portfolio optimization, and codifies research IP; enabling its reuse and preventing its loss. Citrine's customers include Panasonic, LANXESS, and some of the biggest and most respected names in the materials and chemicals industry in Asia, North America, and Europe. For more information visit our website at Citrine.io, or contact us at +1 650-276-7318.