Case Study

INCREASED CUSTOMER RESPONSIVENESS

AI With Domain Knowledge Integration Fills In Data Gaps

AI With Domain Knowledge Integration Fills In Data Gaps

As the pace of product development increases, it is important to be able to respond to customer requirements quickly and competitively. Knowing whether you already have a product meeting the requirements, or if you can make one quickly and cheaply is key. To do this you need data; and it is common to find that not all properties of a material have been measured and recorded due to the time-consuming, expensive nature of testing. It is even more time-consuming to find out if the required property profile can be achieved by a new material within the company's product portfolio envelope.

ARTIFICIAL INTELLIGENCE CAN BE USED TO:

- Predict missing property data in existing products
- Predict the property profile of potential new products
- Optimize portfolio for cost, process efficiency, emissions and inventory

	Young's Modulus	Hardness	Fracture Toughness	Tensile Strength	Cost
Existing Product 1	3GPa	75	?	46MPa	?$/kg
Existing Product 2	?	82	3MPa.m$^{1/2}$	50MPa	?$/kg
Potential Product	?	?	?	?	?$/kg

Citrine's customers are acquiring this capability, so that they can both respond to customers quickly, and bid competitively for contracts

Domain knowledge is essential

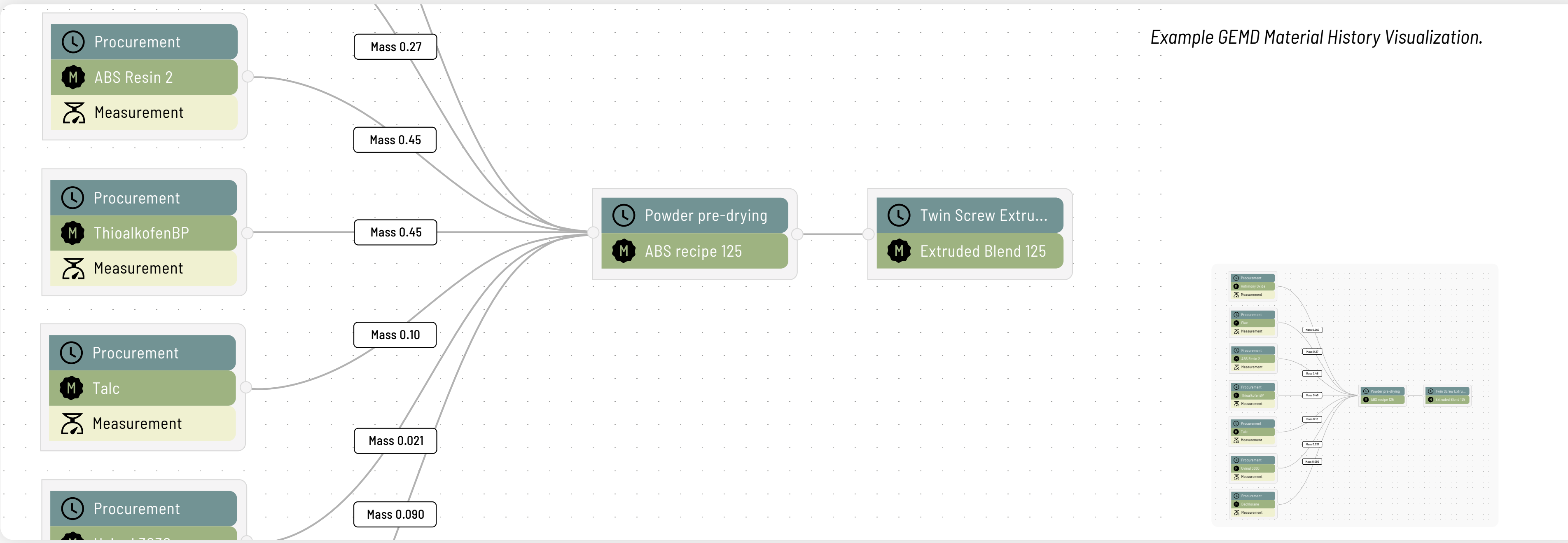
Chemicals and materials producers do not have "Big Data." Testing is costly and time-consuming. Citrine's customer's sometimes start with no data at all, and often start with 30-100 data points. By integrating the domain knowledge of experts, via for example known expressions, AI models can focus on predicting unknowns rather than reinventing the laws of physics. This improves the accuracy of predictions. It is also important to leverage all of the information available. Off-the-shelf AI algorithms read specialized data formats such as chemical formulas and molecular structures as strings of letters and numbers or strange pictures. The Citrine Platform can automatically recognize this type of data and convert it into extra information to be used by the AI model.

Customer example

A global leader in specialty chemicals and plastics whose digital transformation strategy is enabling them to dynamically respond to customer requirements, has worked with Citrine to fill gaps in their data and create an AI driven formulation design workflow.

The Approach

Creating a polymer requires many steps, where different ingredients are added and processed under different conditions. The properties of the end product depend not just on the ingredients added, but the context in which they were added. To be able to predict the properties of the end product all of this context needs to be captured in the data model. Citrine has developed a data model GEMD (Graphical Expression of Materials Data) which captures all of this rich context. The model is flexible enough to work for different material classes, while at the same time enforcing the structure and standards needed to make the data reusable by teams across an enterprise. In this way the customer benefits from synergies between projects across the company. Data entered by one team can be used to improve the model of another team.

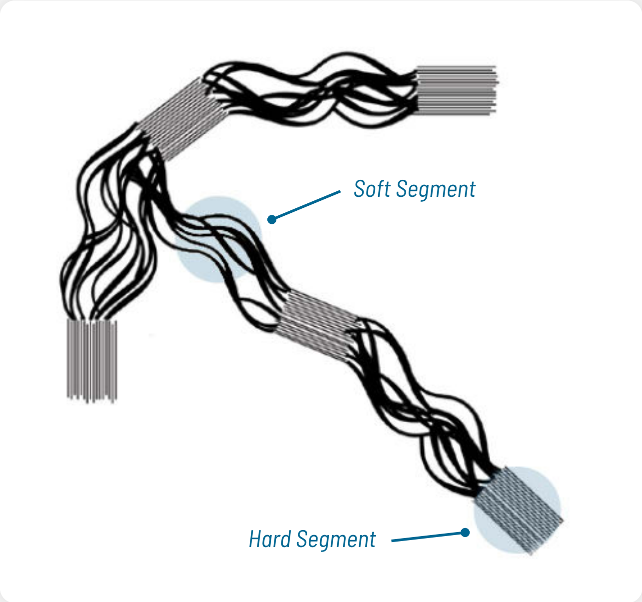


Example GEMD Material History Visualization.

Customer example

Once all of this data is captured, it was important to work with the experts in the customer team to understand the classic processing structure, property relationships. For each target property, there needs to be an understanding of which structural features of the material effect the property, and which aspects of the ingredients and processing are likely to affect this property.

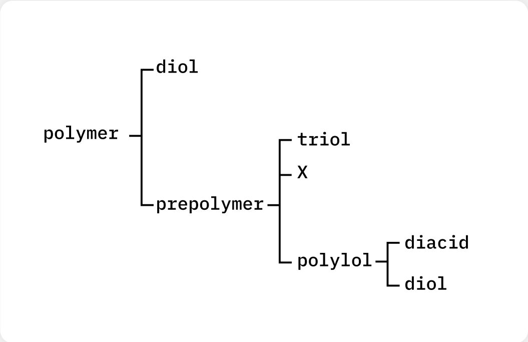
A classic example of this arose for our customer:



In a semi-crystalline polymer, mechanical properties are related to the WT% of soft and hard segments

It was therefore important to know what about the materials history contributed to the hard segments in the material.

The Citrine team was able to "featurize" the material history data into a simplified structure that could be systematically assessed for its contribution to hard segments. As well as looking at polymer structure, other domain knowledge integrated included: the customer's classification system for the polymers, the synthesis context, and information about mass equivalence.



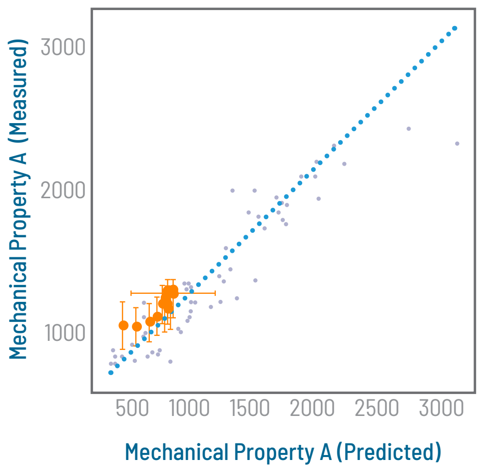
Results

Customer example

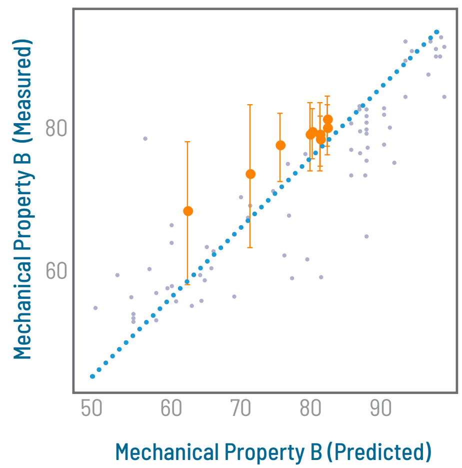
Once this domain knowledge was integrated into the AI model the model was trained on data already available. It could then begin to make predictions for properties of existing materials where data was simply missing and for a whole set of mechanical properties of a hypothetical material not yet synthesized.

The customer chose ~10 samples on which to do a single blind trial. They were tested and compared to model predictions. For 3 properties, the model results had less than 8% error. This already allowed the customer to quickly identify promising recipes for further experimentation and rule out those unlikely to meet the target profile. From there the model was refined further using a process of sequential learning, where data from tests is added to the platform and then used to retrain the model.

Thanks to the flexibility of the model architecture, which is based on the fundamentals of polymer structure, the AI model can now be adapted and used on other polymer types.



Training Error = 0.371
Mechanical Property A



Training Error = 0.518
Property B

Initial model error results (after domain knowledge integration) – compared to validation samples. 0.6 is the threshold below which sequential learning can typically be started.

This model can now be refined using sequential learning.

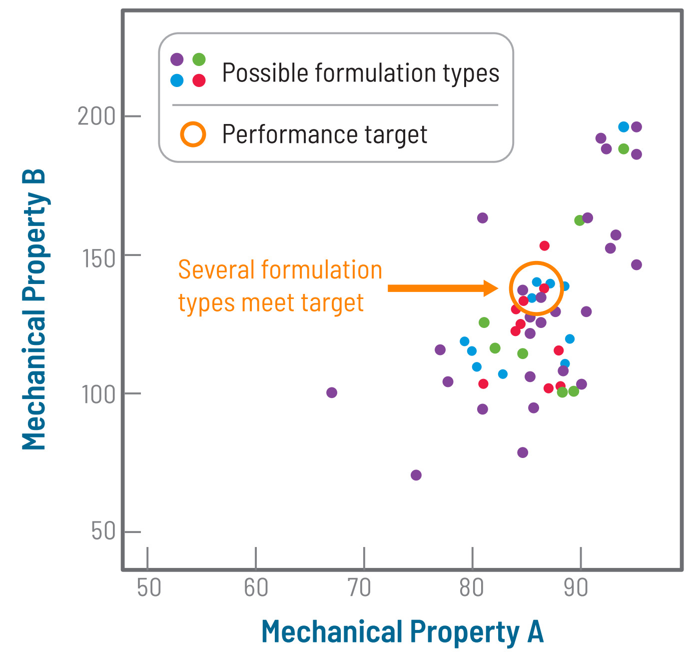
Cost Optimisation

Having the ability to predict the properties of a formulation from basic data, enables expert teams to survey an entire product portfolio for candidates that are likely to meet a customer's requirements. These candidates can then be tested (their data being added into the platform and refining the model automatically). This process reduces the time and therefore cost needed to respond to a customer by reducing the number of experiments needed.

As well as predicting mechanical properties, it is also important to understand production costs. There may be different ways to achieve the required property profile, with different ingredients and processing. By comparing raw material costs, energy use, emissions etc. of the different processing routes, a company can ensure that they make a competitive bid for the contract, whilst retaining appropriate profit margins.

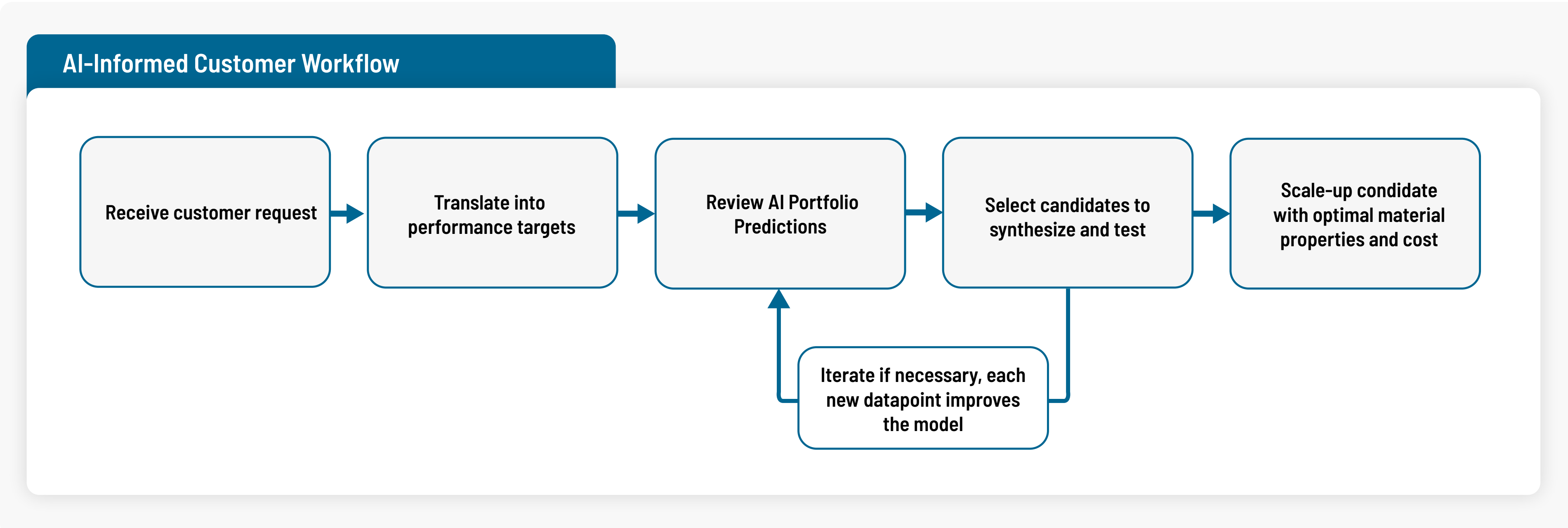
Using a formulation that uses common ingredients may enable them to keep their inventory small and enable bulk buying from suppliers.

Formulations fitting mechanical property criteria



AI-Informed Customer Workflow

Once a model has been validated, it can be used within the customer request workflow.



This workflow doesn't remove the need for laboratory testing or expert teams; instead it focuses them on candidates with the highest probability of success.

Results

Summary

Customer responsiveness is key, in the formulations space, to ensure margins are maintained. AI is being used by specialty chemical companies to act quickly on customer requests. Having an AI platform that is specialised for materials and chemicals, easily integrates domain knowledge and scales as AI is taken up by different teams, is essential.

Arrange an online meeting

Listen to our podcasts

Download more case studies

Subscribe to our newsletter

Citrine Informatics Inc.
2629 Broadway St
Redwood City, CA 94063

citrine.io
info@citrine.io