CITRINE INFORMATICS

White paper

WHAT MAKES A GOOD MATERIALS INFORMATICS SYSTEM?

What makes a good Materials Informatics System?

OVERVIEW

Material Informatics Systems (MIS) add value by accelerating product development, guiding R&D decisions, and capturing and disseminating knowledge across an enterprise. To adopt an MIS, we recommend starting with a couple of proof of concept projects to demonstrate value. However, more value is gained as an MIS is adopted across an organization. This results in shared datasets, AI models, and AI-informed project workflows and decision making. As more assets are created and reused, the benefits compound and value grows exponentially. Before adopting an MIS as part of a digital transformation initiative, it is important to consider what makes a good MIS for your organization.

(It is also worth noting what an MIS is not. It does not coordinate lab personnel, allocate equipment, track samples or log inventory. It is a strategic asset, not a day-to-day work horse.)

Benefits of a Materials Informatics System

1. Create reusable digital assets
2. Visualize and analyze integrated data
3. Discover important processing and composition parameters
4. Optimize new materials or recipes based on multiple property objectives
5. Make strategic decisions about research direction
6. Share knowledge

1. CREATION OF REUSABLE DIGITAL ASSETS

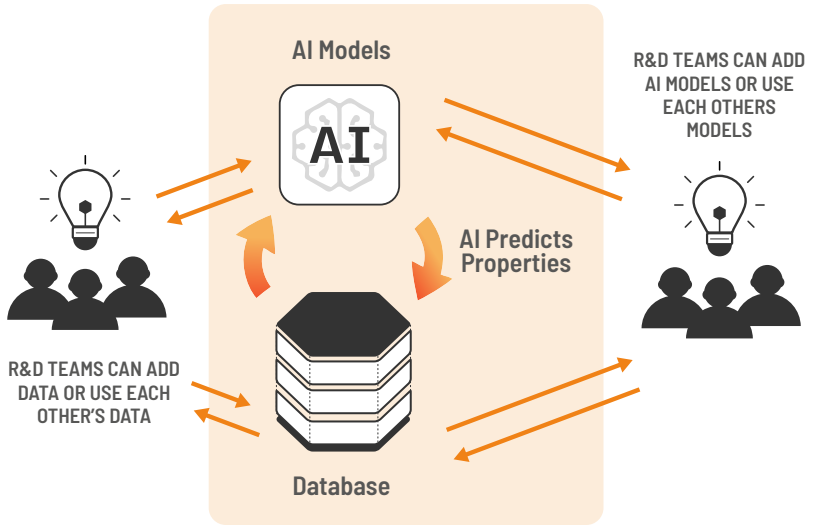
In the materials and chemicals space, collecting new data requires resource-intensive experimentation, so it's unlikely that the AI or machine learning capabilities in an MIS can take advantage of "Big Data".

In order to get the most value out of your existing data, MISs should be able to structure and disseminate data sets across R&D teams, applications engineering teams, and business units.

AN MIS SHOULD HAVE:

- Segregated and encrypted data storage
- An efficient, scalable interface for data ingestion
- A way to share AI-ready data sets across projects and teams
- Tools to establish and standardize data structures
- A data model that captures the full context of each material or chemical

Your data is a valuable reusable asset providing competitive advantage



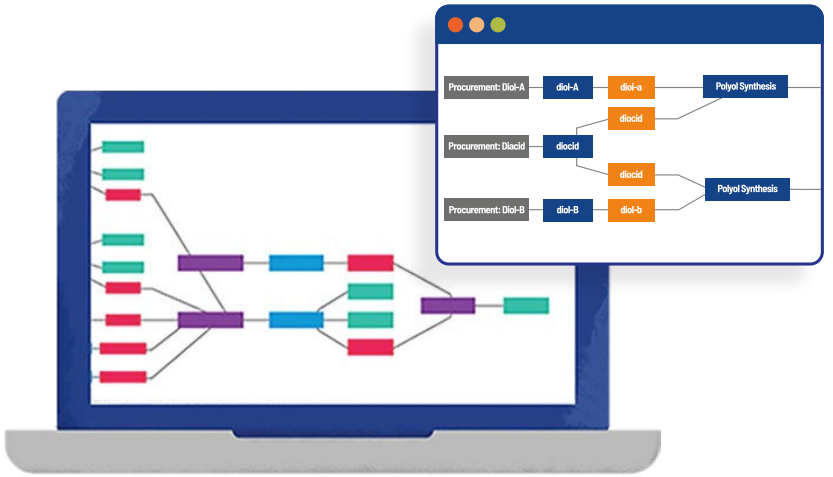
2. VISUALIZE AND ANALYZE INTEGRATED DATA

A material's properties are the result of a whole history of interlinked processing steps and can vary from batch to batch. In order to optimize material performance, it is essential to be able to integrate data from procurement, through processing to characterization. Visualizing this system of complex interactions helps researchers see the big picture and identify relationships within the data.

AN MIS SHOULD HAVE:

- A data model that can capture all the steps to make a material from procurement through to characterization
- A flexible data model that can adapt to new research directions and parameters
- A data model that systematically tracks expected vs actual values of processing parameters
- A way to visualize relationships between data
- A well-designed user-interface

Visualize your data



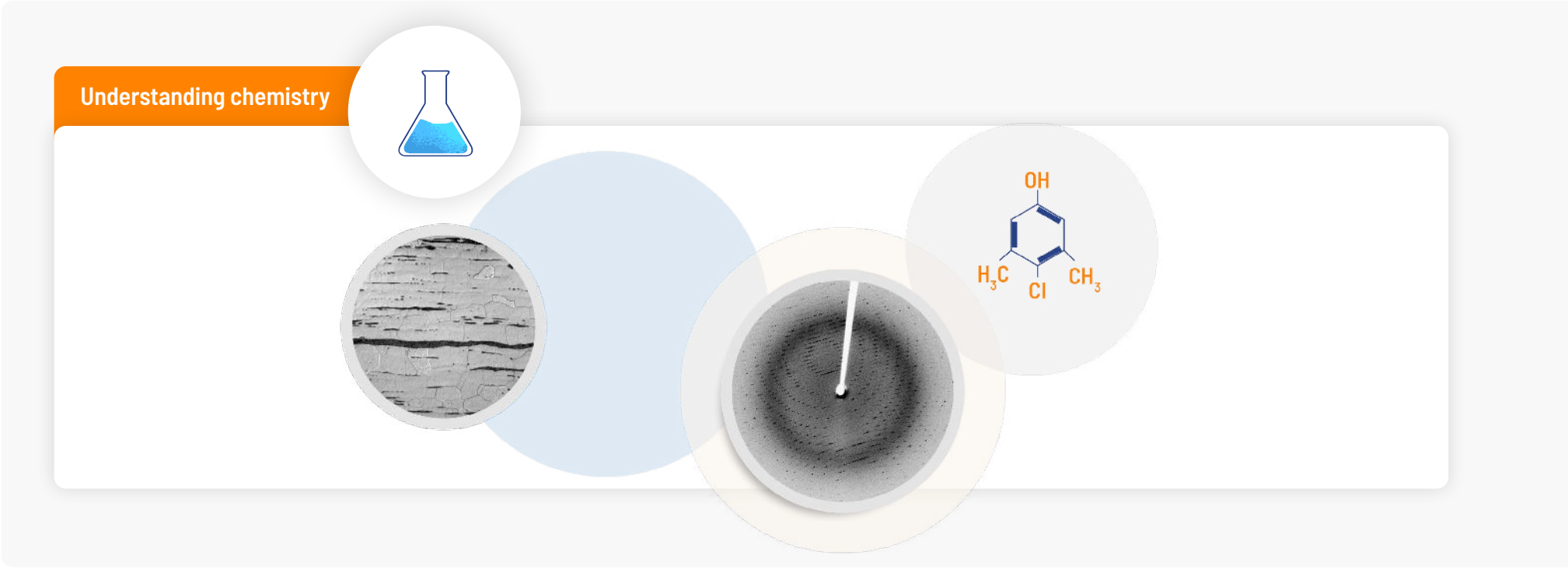
3. DISCOVER IMPORTANT PROCESSING AND COMPOSITION PARAMETERS

AI models take input parameters (e.g. temperature, wt% Carbon, processing time) and predict output properties (e.g. Yield Strength). Some of these parameters will be more important than others in predicting the output property. Understanding the effect of input parameters on outputs is key to both fundamental scientific understanding and achieving scalable production processes.

Materials and chemicals information comes in many forms, such as micrographs, x-ray diffraction images, and chemical formulas. In the case of the micrograph, researchers know which features of the picture, e.g. particle size, have an effect on the property to be modeled. The process of deciding which data is relevant to predicted properties and converting this information to AI-readable data is called featurization. This is a critical step to achieving value from materials informatics.

AN MIS SHOULD HAVE:

- ☑ A library of descriptors that convert materials- and chemical-specific information to data
- ☑ A way to create custom descriptors for new information types
- ☑ A way to easily see what features of the data are having the most effect on the AI model

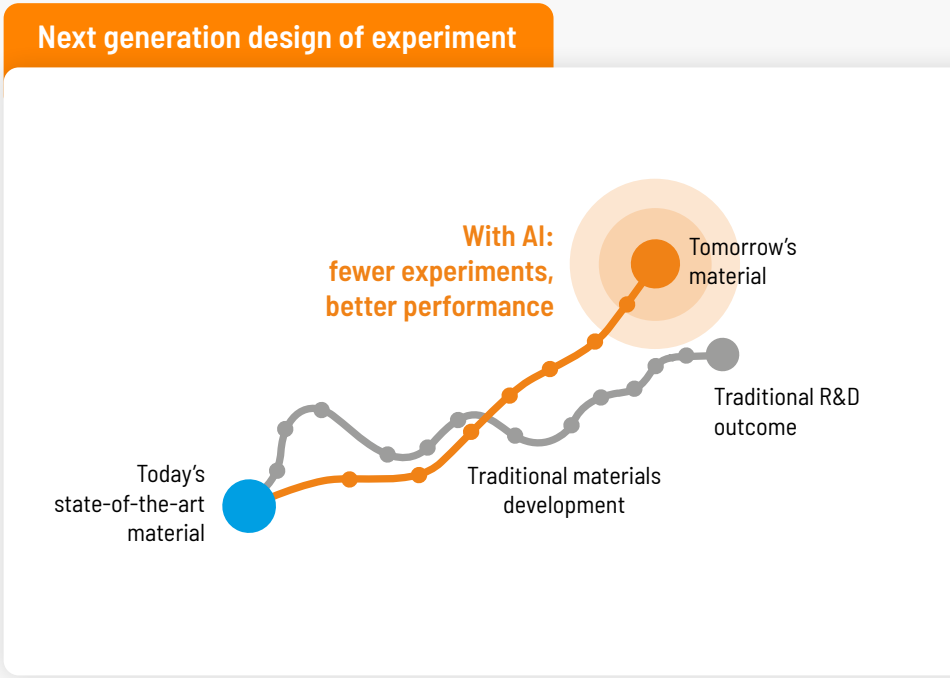


4. OPTIMIZE NEW MATERIALS OR RECIPES BASED ON MULTIPLE PROPERTY OBJECTIVES

Once data is ingested, curated, and featurized, it can be used to train AI models. These models can be used to predict properties of new materials, and, inversely, predict new materials that meet performance characteristics and design, cost, or processing constraints.

AN MIS SHOULD HAVE:

- Materials-specific methods to take advantage of small and sparse data sets
- A workflow that minimizes the number of additional simulations or experiments that are needed to get an accurate AI model
- A way to incorporate scientific or market knowledge into the AI workflow to increase model performance
- A graphical display linking materials candidates to target properties



5. MAKE STRATEGIC DECISIONS ABOUT RESEARCH DIRECTION

Materials and chemicals research is resource intensive and risky. Decision-makers need to know if success is likely before investing additional time and resources into an R&D project. MISs can help with this by incorporating uncertainty estimates into each prediction, and visualizing the likelihood that a given R&D direction will achieve the desired outcomes.

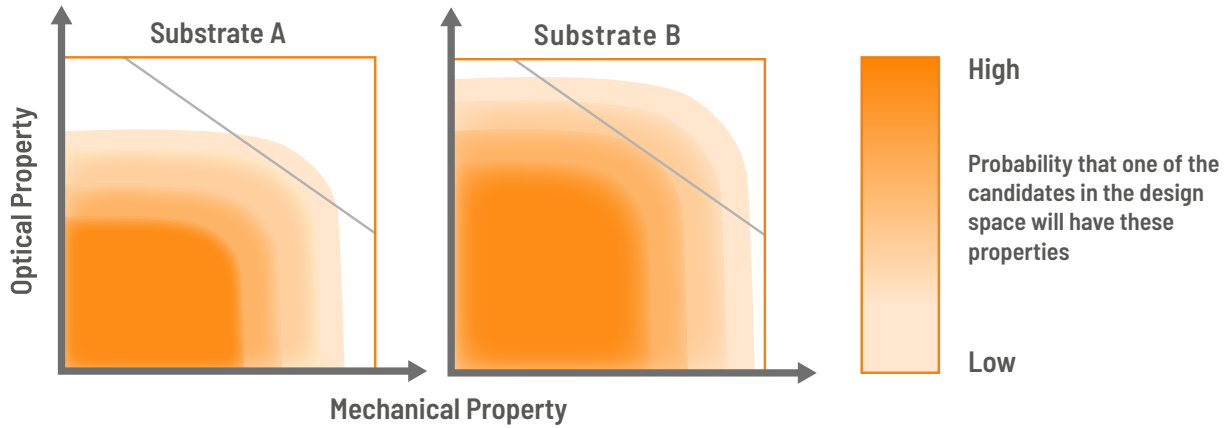
This can help R&D leaders decide:

- whether to continue or kill an R&D project
- if a new expensive ingredient is worth the extra cost
- if there is a good risk balance across the R&D portfolio

AN MIS SHOULD HAVE:

- ☑ Systematic calculation of uncertainty for every prediction
- ☑ A way to sum the probability of success over the design space
- ☑ A way to display this information in an easily digestible format

Compare likelihood of succes in different scenarios



6. KNOWLEDGE SHARING

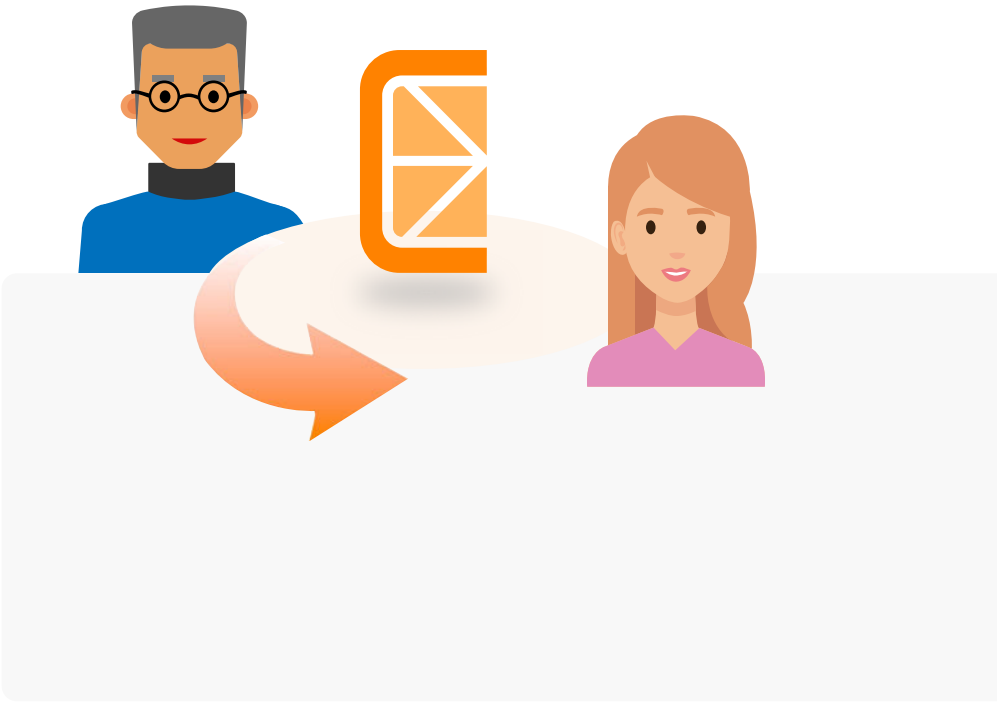
As veteran scientists and engineers retire, and a new generation enters the workforce, it is critical to capture scientific and business domain knowledge and pass that on to help new employees get up to speed quickly.

Researchers need to quickly find out:

- Which processing factors most affect a particular property?
- Do we have any data on….?
- What are the known properties, processes, and structure relationships for the project I'm working on?
- Where did the data come from?
- What was the test standard?
- Which material candidates have already been tested and what were the results?

AN MIS SHOULD ALLOW USERS TO:

- ☑ Find data quickly
- ☑ Identify the provenance of data and what conditions applied when measurements were taken
- ☑ Capture and disseminate domain knowledge in the form of data sets, visualizations, or AI models
- ☑ Access and analyze existing AI Models
- ☑ Assess feature importance in existing models



SUMMARY

A Materials Informatics System is a strategic asset creating competitive advantage by empowering your team to be more agile. To decide which system to invest in, consider:

- Are there good tools to get specialized materials information into the system?
- Does it create reusable digital assets?
- Can it help the team visualize and analyze data?
- Are models easily interpretable, helping researchers to understand how it arrived at its predictions?
- How is the domain knowledge of the team captured and shared?
- Does it provide the data needed for strategic decision making?
- How can your team reuse existing datasets, feature libraries, models, and visualizations?

Contact Citrine Informatics to find out how the Citrine Platform performs against these criteria.

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

citrine.io
info@citrine.io