

CITRINE

INFORMATICS

USING AI TO: REDUCE RAW MATERIAL COSTS AND LOWER THE RISK OF FAILED FINAL COMPLIANCE TESTS

How materials and chemicals teams can cut wasted development effort by screening smarter before the most expensive tests begin

EXECUTIVE SUMMARY

Across materials and chemicals, development teams are under pressure to do two things at once: **reduce formulation cost** and **maintain compliance**. In many workflows, those goals collide late in the process, when teams have already invested heavily in candidate formulations only to discover that a final test fails. That failure may come in a large-scale fire test, a toxicology screen, or another compliance-critical validation step that is expensive, slow, and difficult to repeat.

This is where AI changes the economics of product development.

Instead of treating final compliance testing as the main discovery mechanism, teams can use AI to learn from earlier-stage data, rank candidate formulations more effectively, and screen out weak options before they consume expensive testing capacity. The result is not lower standards – it is a better path to meeting them.

TWO ILLUSTRATIVE CASE STUDIES

Polyurethane Foams (Huntsman Building Solutions)

Used the Citrine Platform to reduce reliance on large-scale fire testing, shorten development time, and deliver a regulation-compliant foam. Starter formulation creation time reduced from **a few months to less than 1 day**.

Agri-Chemical Formulations Toxicity Screening

Citrine developed a toxicity model in **4 weeks** with **82% pass-fail accuracy**, enabling teams to screen out likely-toxic formulations before lengthy product development and expensive animal testing occurred.

i For R&D leaders and managers, the implication is practical: the most valuable use of AI is often not broad automation – it is reducing wasted effort before the costliest compliance decisions are made.

THE REAL PROBLEM IS NOT JUST COST – IT IS LATE-STAGE FAILURE RISK

Most materials and chemicals organizations already know how to optimize one variable at a time. They know how to reduce ingredient cost. They know how to improve performance. They know how to run compliance tests. What makes development slow and expensive is having to **optimize all three under real-world constraints**.

COST SUBSTITUTION RISK

A lower-cost formulation may change flame spread, smoke development, toxicity, durability, or other regulated outcomes.

SUPPLY RESILIENCE RISK

A substitution that improves supply resilience may introduce new uncertainty in final validation.

SEARCH SPACE COMPLEXITY

Teams work in a space too large to explore exhaustively, while the most definitive tests are also the most expensive to run.

That creates a familiar pattern. Teams make their best early decisions with incomplete evidence, narrow the field through trial and error, and then commit significant time and budget to a final compliance test. If that test fails, the cost is not only the test itself – it is the accumulated delay, reformulation work, and opportunity cost behind it.

WHY TRADITIONAL WORKFLOWS STRUGGLE

In conventional development, expensive compliance tests often arrive **too late to guide exploration efficiently**. By the time a candidate reaches a large-scale fire test or a final animal toxicity screen, the team has already made many decisions about ingredients, ratios, and processing variables.

That sequence makes sense operationally, but it is a poor fit for complex formulation problems. It allows too many weak candidates to survive too long, and it gives teams limited feedback about which early signals actually predict late-stage pass or fail outcomes.

⚠️ This is especially problematic when the final test is external, capacity-constrained, ethically sensitive, or costly enough that only a small number of candidates can be evaluated. In those settings, each failed final test is a **material business event**, not a routine lab result.

WHAT AI CHANGES

AI does not replace expert scientists or final compliance testing. It helps teams decide **what deserves to reach that stage.**



START WITH EXISTING DATA

Formulation history, ingredient properties, process variables, and results from earlier-stage tests form the foundation of the AI workflow.



PREDICT MULTI-OBJECTIVE PERFORMANCE

The model predicts which candidate formulations are most likely to meet cost, technical performance, and compliance-related outcomes simultaneously.



INCORPORATE DOMAIN KNOWLEDGE

Constraints, ingredient roles, featurization, and known chemistry relationships are added to enrich the model's predictive capability.



SCREEN AND PRIORITIZE

Candidates likely to fail are screened out early. Stronger options are prioritized. Expensive compliance tests become targeted validation, not broad search.

That is the shift: final tests still matter, but they stop carrying so much avoidable discovery burden.

EXAMPLE 1: POLYURETHANE FOAMS AND COSTLY FIRE TESTING

THE CHALLENGE

Huntsman Building Solutions was developing closed-cell polyurethane foam where meeting flammability requirements was essential. Large-scale fire tests such as **ASTM E84** were complex, expensive, time-consuming, and conducted off site by an independent tester. Huntsman wanted to reduce the number of those tests needed to create a compliant product.

THE AI APPROACH

Because there were too few large-scale fire-test data points alone to build a strong model, the workflow incorporated **small-scale fire-test data** as well. Huntsman then used iterative modeling and experimentation, improving the featurization of those small-scale tests as new results came in.

THE OUTCOME

A compliant formulation with improved fire resistance that also retained sprayability and hit physical performance targets – with reduced reliance on full-scale fire testing and shortened development time.

<1 DAY

STARTER FORMULATION

Reduced from a few months to less than 1 day for PU foam development



FIRE TEST RELIANCE

Reduced reliance on full-scale ASTM E84 fire testing through small-scale signal modeling

- ✔ This example shows that when the final compliance test is expensive and infrequent, AI can still create value by learning from cheaper, earlier signals and using them to improve experiment prioritization.

EXAMPLE 2: SCREENING OUT TOXIC FORMULATIONS BEFORE FINAL ANIMAL TESTING

CASE AT A GLANCE

A customer developing complex multi-phase chemical formulations needed to reduce the number of formulations that reached final animal toxicity testing only to fail.

Data Available: 36 formulations, expanded to 108 rows through dilution testing

Model Development Time: 4 weeks

Cross-Validation Accuracy: 82% pass-fail prediction

Hold-Out Set Performance: 5 out of 6 correct qualitative toxicity results and quantitative toxicity scores

WHY THIS MATTERS

The challenge was to predict one aspect of final-product toxicity based on ingredients, formulation structure, and dilution data. The model could be used to **screen out formulations likely to be toxic** before lengthy product development and expensive testing occurred.

This example broadens the business case for AI. The value is not limited to speeding up formulation iteration. It is also about reducing the risk that teams will carry weak candidates all the way to a final compliance gate that is costly, slow, and undesirable to repeat.

- ✔ A global specialty chemicals company built a toxicity model with 82% accuracy in 4 weeks, enabling toxic recipes to be screened out before expensive development and reducing animal testing.



82%

PASS-FAIL ACCURACY

Cross-validation accuracy in predicting toxicity test outcomes



5/6

HOLD-OUT PREDICTIONS

Correct qualitative and quantitative toxicity predictions on unseen data



4 WKS

TIME TO MODEL

From project start to deployable toxicity screening model

A GENERAL PATTERN FOR MATERIALS AND CHEMICALS TEAMS

Taken together, these cases show a **repeatable pattern** across materials and chemicals product development.

FINAL COMPLIANCE TESTS ARE NECESSARY BUT INSUFFICIENT AS SEARCH ENGINES

The final compliance test is important but too expensive to serve as the main mechanism for exploring the formulation space. Its role should be targeted validation, not broad discovery.

EARLIER-STAGE DATA CONTAINS UNDERUTILIZED SIGNAL


Earlier-stage data usually contains signal that can be made more useful through good featurization and modeling – even when that data is sparse or imperfect.

AI REDUCES WASTED EFFORT BY IMPROVING CANDIDATE SELECTION

AI can reduce wasted effort by improving which candidates move forward, ensuring that constrained testing capacity is directed at the most promising options.

THE OUTCOME IS LOWER RISK, NOT JUST FASTER DEVELOPMENT

The outcome is not only faster development – it is lower risk of late-stage failure after substantial investment has already been made.

 Cost optimization tends to increase the number of candidate substitutions and reformulation paths that must be considered. Without a better screening mechanism, that broader search space can **increase compliance risk** rather than reduce it.

WHAT THIS MEANS FOR R&D MANAGERS

For an R&D manager, the practical question is not whether AI is interesting. It is **where AI can lower the cost of decision-making now.**

THE STARTING POINTS SHARE THREE CHARACTERISTICS



COSTLY FINAL TEST

A high-stakes compliance gate that is expensive, slow, or ethically constrained



EARLIER-STAGE SIGNALS

Existing data from smaller, cheaper tests that may predict final outcomes



COST OR SPEED PRESSURE

Active pressure to reduce ingredient cost or reformulate quickly

That could be a large-scale fire test, a final toxicology screen, a durability qualification, or another high-stakes validation step. In each case, the value of AI is similar: it gives scientists a better basis for deciding what to test next, with the messy data they already have.

CONCLUSION

Across materials and chemicals, some of the most expensive development failures happen at the end of the process, when a formulation reaches a final compliance test and does not pass. That is the moment AI can help teams avoid earlier.

The opportunity is straightforward: reduce ingredient costs without increasing late-stage risk, and use early signals to screen smarter before the most expensive tests begin.

The PU foam fire-testing example shows how this works when large-scale flammability testing is the gate. The toxicology case shows the same logic in a formulations workflow where final animal testing is the gate. In both cases, the value comes from reducing wasted development effort before the last and costliest decision.

THE CITRINE PLATFORM

Trusted by 10 of the top 20 specialty chemicals companies, with hundreds of millions in proven value for our customers.

380

PROJECTS/MONTH

Active projects worked each month

550

AI MODELS/MONTH

AI models deployed each month

70K

SUGGESTIONS/MONTH

Experiment suggestions generated each month

6.5

DAYS TO FIRST SUGGESTION

Days to first experiment suggestions (Median since 2021)

25

WEEKS TO FIRST WIN

Average weeks to first technical win (10 customers surveyed)

1600%

EXAMPLE FIRST YEAR ROI

Specialty chemical company

READY TO REDUCE THE RISK OF YOUR INNOVATION PIPELINE?

[Request a meeting](#)