



IRI 2026

INNOVATORS SUMMIT

May 6-8 / Charlotte, NC





AI in R&D and Innovation

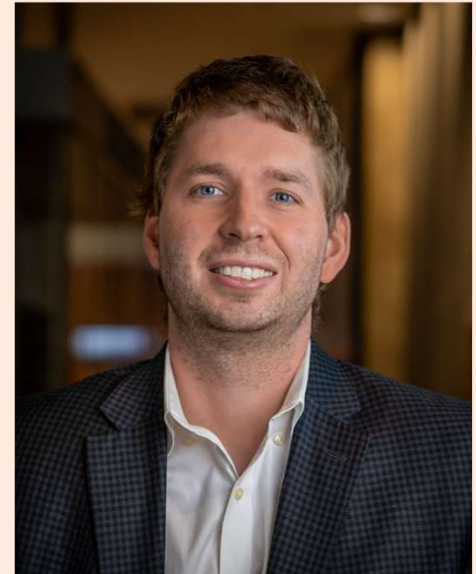
Presented by Noel Hollingsworth | CEO & Co-Founder | Uncountable

IRI Innovators Summit
Charlotte, NC
5/6/2026

CONFIDENTIAL: This document contains proprietary information. Unauthorized sharing or distribution is strictly prohibited.

Meet The Speaker: Noel Hollingsworth

CEO & Co-Founder | Uncountable



CEO & Co-Founder

Noel Hollingsworth

Noel Hollingsworth is the **CEO and Co-Founder** at **Uncountable**. In his role, he works closely with Uncountable's customers to implement next-generation data management systems.

Prior to his work at Uncountable, Noel **led data teams** at startups and was awarded **Forbes 30 under 30** for his **work with machine learning and artificial intelligence**.

About Uncountable



Originally Founded in 2016

- **Offices:**
 - San Francisco, CA (HQ)
 - New York City, NY
 - Munich, Germany
 - London, UK
 - Singapore



>150 Customers, >1000 Labs

- **Enterprise R&D Teams:**
 - Paints & Coatings
 - Cosmetics
 - Advanced Materials
 - Food & Beverage
 - Pharma & BioTech
 - CPG
 - Agriculture



#1 AI Platform for End- to-End Product Development

- **One Source of Truth:**
 - Centralized data management, capture, & access across all teams
 - Reduces new product dev. timelines & iterations
 - Predictive analytics & formulation suggestions
 - Connected data across systems and databases



Proven Domain Expertise

- **Founded by Experts**
 - Began as a data science company helping Fortune 500 materials companies accelerate development of new projects
 - Leveraged learnings to develop a platform focused on structured data to enable proper AI-driven insights

Multinational Enterprises Across a Variety of Industries

Uncountable Proudly Supports Clients

Beiersdorf

Mondelēz
International

tesa

DUPONT

Jika

BEHR

Lubrizol

syngenta®

REPSOL

SunChemical
a member of the DIC group
CFC
Color & Comfort

CLARIANT^E

AGC

SCG

INDORAMA
VENTURES

DOW

syngenta

ZEON

TotalEnergies

CooperStandard

ROGERS
CORPORATION

TRINSEO

KRATON

CABOT

AirBoss
Rubber Solutions

NAZDAR

3D SYSTEMS

MEGGLE

LESAFFRE

BlueNalu

JOTUN

Alianza
Team

6K

Shurtape
TRUE TO YOUR WORK

coim

SOPREMA

cmbllu

remmers

GEON[®]
Performance Solutions

VALTRIS
SPECIALTY CHEMICALS

IRI 2026
INNOVATORS
SUMMIT

M.

Agenda

- 01 **Introduction**
- 02 The Current Landscape: AI in R&D
- 03 Pitfalls of Common Data Systems & Sources
- 04 Setting Expectations & Considerations Before AI-Driven R&D
- 05 Creating a Roadmap for AI-Driven R&D
- 06 The Benefits of a Unified R&D Lab Data Infrastructure
- 07 Q&A





The AI Imperative

- **The Need to Innovate The Process of Innovating - GO FASTER**
 - R&D is undergoing a major transformation: From trial-and-error to data-driven design
 - AI enables faster innovation cycles, deeper insights, and cost savings.
- **R&D faces challenges across segments worldwide - SOLVE HARDER PROBLEMS**
 - Global Competition
 - Regulatory & Compliance Complexity
 - Sustainability Mandates & Ingredient Sourcing
 - Data Silos & Poor Data Infrastructure
 - Complex Consumer & Market Demands

IRI 2026

**INNOVATORS
SUMMIT**



Leveraging AI in R&D: Achieving a Competitive Advantage

Future Prospects & Trends in AI-Driven R&D

- From Assistive to Autonomous R&D
- Foundation Models & LLMs in Scientific Workflows
- AI-Augmented Human Expertise
- Scalable Institutional Knowledge Capture & Use
- Sustainability-Driven Innovation
- AI for Multiscale & Multimodal Science

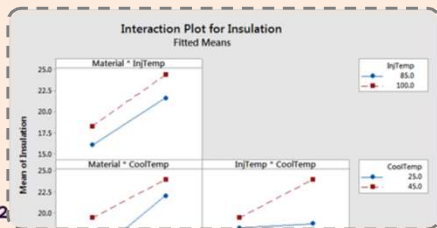
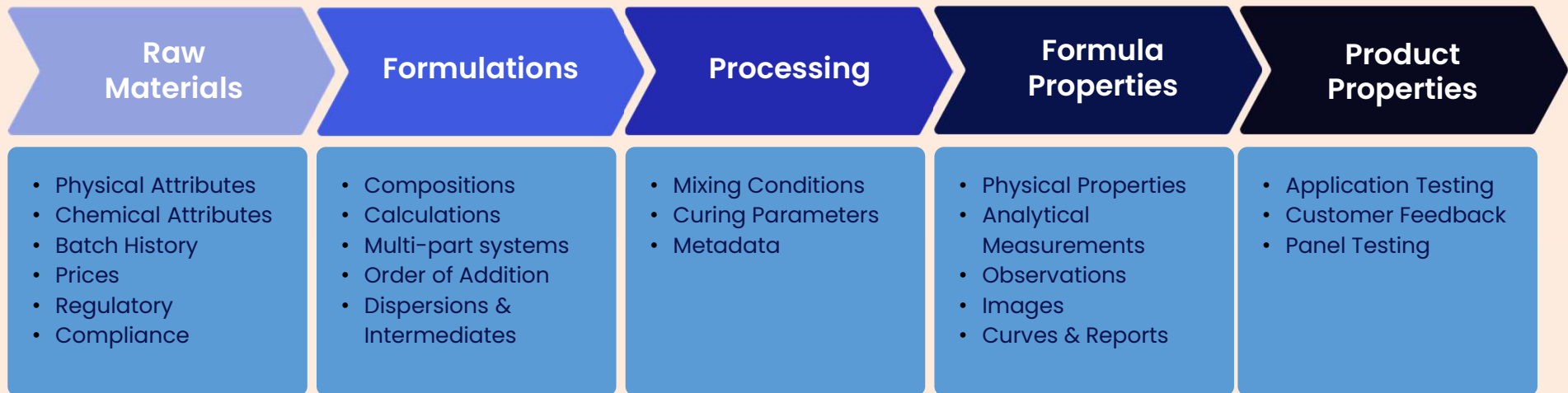
Ways AI Can Enhance R&D Performance & Competitiveness

- **Short-Term**
 - Accelerated Data Analysis & Pattern Recognition
 - Reduced Time-to-Insight
 - Improved Decision-Making
 - Workflow Automation
 - **Cost Optimization in Experiments**
- **Long-Term**
 - Faster Innovation Cycles
 - Improved Product Success Rates
 - Knowledge Retention & Reuse
 - **Competitive Differentiation**
 - **Scalable Collaboration Across Functions & Sites**

Decentralized, Unstructured, & Fragmented Data

What stands in our way?

Data is collected & sits independently across the different teams and systems used throughout the entire R&D value chain...



Available Factorial Designs (with Resolution)

Runs	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1														
2														
3														
4														
5														
6														
7														
8														
9														
10														
11														
12														
13														
14														
15														

Available Resolutions III (Plackett-Burman Designs)

Factors	Runs	Factors	Runs	Factors	Runs
2-7	12,28,24,28,...48	29-23	24,28,32,36,...48	36-39	48,44,48
8-11	12,28,24,32,...48	24-27	28,32,36,40,44,48	40-43	44,48
12-15	28,24,28,36,...48	29-31	32,36,40,44,48	44-47	48
16-19	28,24,28,32,...48	32-35	36,40,44,48		

NAME	Parts	Batch Weight	Parts	Batch Weight	Parts	Batch Weight	Parts	Batch Weight	Parts	Batch Weight
Polymers A	21	21	14	14	6	6	6	6	1	1
Polymers B	16	16	16	16	16	16	16	16	16	16
Carbon Black C	0	0	11	11	0	0	25	25	10	10
Carbon Black D	0	0	0	0	17	17	0	0	0	0
Plasticizer	30	30	11	11	7	7	11	11	11	11
Processing Aid	0	0	4	4	0	0	0	0	4	4
Adhesive										
Colorant/Ink										
Slits	21	21	14	14	11	11	21	21	14	14
Coatings	1	1	1	1	1	1	1	1	1	1
Coating Agent										
Coil	1	1	1	1	1	1	1	1	1	1



Agenda

- 01 Introduction
- 02 The Current Landscape: AI in R&D
- 03 Pitfalls of Common Data Systems & Mindset**
- 04 Setting Expectations & Considerations Before AI-Driven R&D
- 05 Creating a Roadmap for AI-Driven R&D
- 06 The Benefits of a Unified R&D Lab Data Infrastructure
- 07 Q&A



Common Data Systems in R&D Labs

LIMS, ELNs, ERPs, & Misc. Internal Databases



Lab Info Management Systems "LIMS"

- Sample management
 - Metadata
- Output Capture
- Task management



Electronic Laboratory Notebooks "ELNs"

- Experimental capture in real time / Collaborative
- Needs "spreadsheet" support
 - *Formulation/analysis*
- Unstructured



Enterprise Resource Planning "ERPs"

- Inventory Systems
- Equipment Management
- BOMs
- Information Store / System of Record



Other Internal Databases

- Inventory Systems
- Equipment Management
- BOMs
- Information Store / System of Record

Unstructured vs. Structured Data:

	Unstructured	vs.	Structured
Examples	<ul style="list-style-type: none">• Spreadsheets• Word Documents• PDFs• Lab Journals/ELNs• Shared Drive		<ul style="list-style-type: none">• Databases• LIMS• Inventory Systems• Uncountable
Advantages	<ul style="list-style-type: none">• Free• Unrestricted entry of information• Known/second nature “habitual”		<ul style="list-style-type: none">• Instant access to specific information/data• Shareable & scalable information• Intelligent insights & reporting
Disadvantages	<ul style="list-style-type: none">• Limited scope & scalability for application of info• Ctrl+F keyword searching• Limited collaboration• Inability to innovate efficiently & at market-rate		<ul style="list-style-type: none">• Requires intentional/deliberate entry of information• Change management• Migration of historical data into new system• Disciplined use

IRI 2026

**INNOVATORS
SUMMIT**



Importance of Structuring Lab Data for AI:

Example of Brookfield Viscosity

- **Standard Way Data Gets Recorded In Spreadsheets and Notebooks:**
 - Viscosity, 7D = 3000
 - Brookfield Visc. Sp #4 = 5500
 - BV, ON = 1800
- **Best Practices for Structuring Lab Data for AI:**
 - Brookfield Viscosity = 5000
 - Liquid Aging Time + Temperature: 7D at 23°C
 - Spindle #4
 - RPM: 150
 - Test Temperature: °23
 - Exact temperature and time
 - Machine SN, Operator

Top 3 Problems Deploying AI Without Structured Data

Why Excel & Unstructured Data System Are Insufficient

1

Volume of Data

- A small data set with the best AI model in the world is **worse** than both expert scientists and simpler AI models applied to **“big data”**
 - The most important aspect of any AI model is its **underlying data** - both **size and cleanliness**

2

Relevancy to Problems

- Desire to **squeeze square peg in round hole** - When we do have some data, we must apply AI, even if it's not a fit
 - AI is **not a fit for all** use cases!

3

Scientist Trust

- Desire to be an AI first company without gathering appropriate data results in scientist trust being lost
 - AI ends up being applied to **unfit projects or only high priority projects** with substantial risk of failure. Issues causes team to lose faith in process
 - **Sufficient Data is important, but not the only prerequisite**



Can LLMs structure data for us?

- A little bit, but not enough
- An LLM is really good at understanding that “Sp. 4” refers to Spindle 4
- An LLM is not good at knowing which spindle number was used, when the researcher didn’t include that information
- An LLM will overconfidently infer generalizations about your data, adding noise

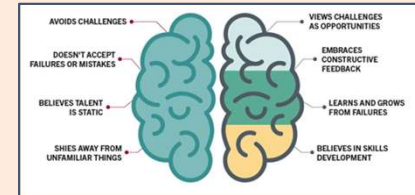


Can LLMs learn from this data for us, without structuring it?

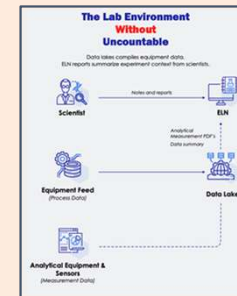
- Maybe at a high level, but not more than a scientist
- In practice, even with a set of well structured data, R&D teams have data that is high dimensionality, low in number of data points
 - Traditional machine learning models want number of data points equal to the number of inputs squared to learn the full space well
- A capable LLM harness will hand off to a specialized model built for this space, such as a Gaussian Process

What about process problems?

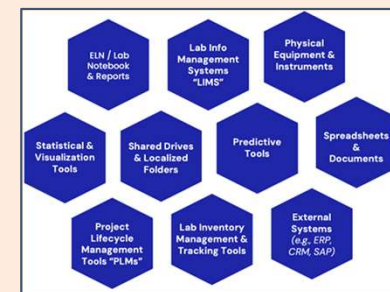
1. Academic vs Enterprise Mindset



1. Data Model coming last



1. Isolated Experimental Capture



Academic vs Industrial Mindset

Academic

Focus

- Massive changes are the only thing that matters
- “superstar” scientist model – believe gains come from them operating however they want
- Allow scientists to still have independence on tool to choose



Enterprise

Focus

- 5-20% improvement on existing experiments may be valuable
- Scientists still important but emphasis on infrastructure and team rather than single person
- OK mandating what scientists will use as a tool

Data Model Coming Last

- Today's Landscape – Many point systems that often work well for their area of expertise
- Data Model is introduced after – we have these systems, let's throw them in a data lake
- What data model should we use? Whatever is imposed by systems (which may not be consistent)

Quote from Peter Norvig (Google Research Director)

"More data beats clever algorithms, but better data beats more data."

IRI 2026

**INNOVATORS
SUMMIT**



Isolated Experiment Capture

- Corollary to previous point – often systems are never joined in the first place
- Lab notebooks not ingested into data lake
- LIMS systems that capture results but miss scientific context of what was accomplished
- Equipment sensors that capture in-depth data, but aren't tied to experiments

Discussion Topic – what problems are you facing?

- What data problems are your teams facing?
- How much of these are software problems vs process problems?
- If you could design the perfect software solution for your teams, would they use it tomorrow?

Agenda

- 01 Introduction
- 02 The Current Landscape: AI in R&D
- 03 Pitfalls of Common Data Systems & Sources
- 04 Setting Expectations & Considerations Before AI-Driven R&D**
- 05 Creating a Roadmap for AI-Driven R&D
- 06 The Benefits of a Unified R&D Lab Data Infrastructure
- 07 Q&A



How can we solve these problems?

- **Accept that not all use cases will find a system**
 - Try for interoperability where we can
 - Better to have 80% of data integrated, than to capture 100% of data in “paper on glass” format
- **Place the data model first**
 - Ensure that when you do capture data, it aligns to a data model created in-house
 - Scientific model of what happens in an experiment– not just can it fit into our data warehouse
 - 80% Data captured, joined and harmonized to a preexisting data model will allow for better team to team learnings when each team creating their own standards
- **Change Management**
 - Most important piece – reason many projects fail

IRI 2026

**INNOVATORS
SUMMIT**



Tips for Team Readiness & Change Management

Adapt workflows gradually:

Update processes step by step to reduce resistance.

Build trust early:

Show quick wins to gain confidence in the new system.

Address technical debt :

Clean up legacy issues before scaling.

Define a single source of truth:

Ensure everyone relies on the same data.

Drive collective use & encourage adoption, not just buy-in:

True success comes when the system is fully embraced in daily work.



Low Code Configuration and Change Management



Example Software Partnership

Configuration

Platform configured to meet needs of the scientists and technicians so there is no desire to use spreadsheets or legacy systems. Vendor should focus on understanding the R&D organization and delivering a product that exceeds expectations.

Trainings from a global team

Training users on their specific use cases assures that users pick it up faster and aren't left with only a help-site to learn key features and functionality.

Encouraging Usage

Lab Requests

It's simple, if lab requests are created and managed in the software system everyone will use it frequently.

Analysis + Visualizations

By discovering critical insights and easy ways to analyze the data in system, rather than exporting to other tools, scientists are encouraged to use the core software platform as a single source of data entry.



Considerations: Setting The Right Expectations

Too Big of a Search Space

- 100s of ingredients, but limited data points
- Either from collection, cleanliness, or standardization

Moonshot Objectives

- What are you trying to achieve in this project vs long term goals
- What are more reasonable targets that would allow you to claim “progress”

Perception of Perfection

- Why would model suggest such a thing?
- Why isn't model more accurate?
- Can it model pictures of exposure ratings?

IRI 2026

**INNOVATORS
SUMMIT**



Agenda

- 01 Introduction
- 02 The Current Landscape: AI in R&D
- 03 Pitfalls of Common Data Systems & Sources
- 04 Setting Expectations & Considerations Before AI-Driven R&D
- 05 Creating a Roadmap for AI-Driven R&D**
- 06 The Benefits of a Unified R&D Lab Data Infrastructure
- 07 Q&A



Evaluating AI Readiness for R&D Applications & Workflows

3 Key Takeaways from Peter Norvig (*Google Research Director*):

- 1. Data drives progress:**
It's not about better algorithms; it's about having more (and cleaner) data.
- 1. Quality matters most:**
More data beats smart algorithms, but better data beats more data.
- 1. Keep it simple:**
Simple models with lots of data often outperform complex models trained on limited data.

What This Means for R&D Teams:

- 1. Prioritize Data Collection:**
Focus on gathering more diverse, high-quality data.
- 2. Emphasize Data Quality:**
Ensure data is "better" — clean, structured, and relevant to R&D goals.
- 3. Match Models to Data:**
Choose models that align with the data available, even if they are simpler.

Long Term Data Maturity Model for R&D Orgs

Document Data via Pen & Paper

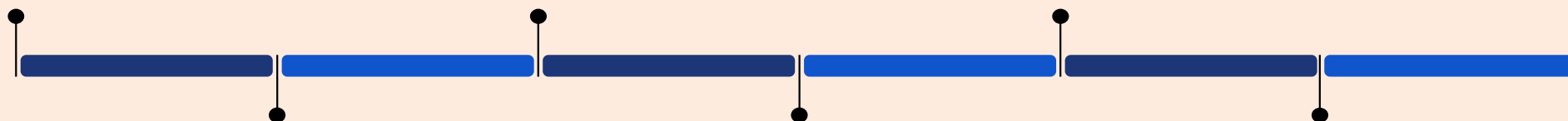
Physical Lab Notebooks/Journals

Centralize Digital Data Storage

Sharepoint, Electronic Lab Notebooks

Use Organizational Data for AI Models

Uncountable's "Unified Laboratory Informatics Platform"



Collect Data Digitally

Excel, Word Documents

Structure Organization-Wide Data

Uncountable's data collection platform

Regular AI-Driven Experimentation

AI tech, org-wide discipline, & standard setting

IRI 2026

**INNOVATORS
SUMMIT**



Defining & Creating an AI Roadmap

BEFORE

Preparation

1. Ensure structured data system is in place/implemented
2. Verify all scientist work is being captured in a way fit for AI
 - All data points and all aspects of data
 - Example: Viscosity centipoise, temperature, spindle, rpm...
3. Reference in-house expertise to understand and validate vendor & partner claims

DURING

Deployment

1. Identify appropriate targets for AI - Example Criteria:
 - Large Amounts of Data
 - Known Success Criteria
 - Consistent Output Results
1. Ensure AI is embedded into daily workflows
 - Not judged off success in a project where majority of results are out of scientists control

AFTER

Maintenance

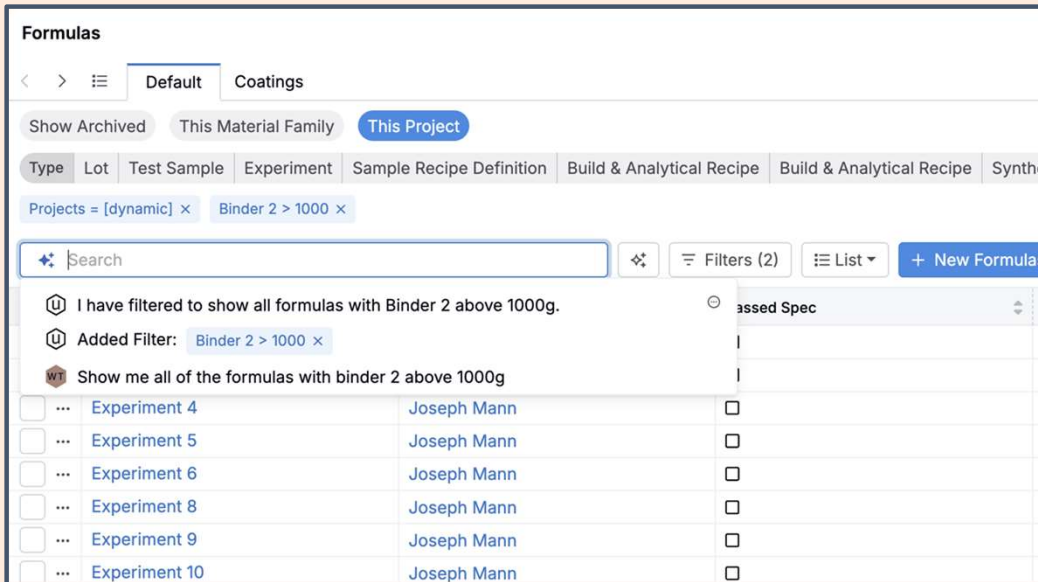
1. Identify areas where data capture is insufficient
 - 1. Deploy systems and/or recurring procedures to collect data

Agenda

- 01 Introduction
- 02 The Current Landscape: AI in R&D
- 03 Considerations Before and After Implementing AI
- 04 Setting Expectations & Considerations Before AI-Driven R&D
- 05 Creating a Roadmap for AI-Driven R&D
- 06 The Benefits of a Unified R&D Lab Data Infrastructure**
- 07 Q&A



Advantages for LLM Usage with Structured Data



The screenshot shows a web application interface for 'Formulas'. At the top, there are tabs for 'Default' and 'Coatings'. Below that, there are buttons for 'Show Archived', 'This Material Family', and 'This Project'. A navigation bar contains tabs for 'Type', 'Lot', 'Test Sample', 'Experiment', 'Sample Recipe Definition', 'Build & Analytical Recipe', and 'Synthe'. There are also filter tags: 'Projects = [dynamic] x' and 'Binder 2 > 1000 x'. A search bar is present with a search icon and a 'Search' label. To the right of the search bar are buttons for 'Filters (2)', 'List', and '+ New Formulas'. Below the search bar, there are three chat-like messages: 'I have filtered to show all formulas with Binder 2 above 1000g.', 'Added Filter: Binder 2 > 1000 x', and 'Show me all of the formulas with binder 2 above 1000g'. Below these messages is a table with columns for checkboxes, experiment names, and names.

...	Experiment	Name	...
<input type="checkbox"/>	Experiment 4	Joseph Mann	<input type="checkbox"/>
<input type="checkbox"/>	Experiment 5	Joseph Mann	<input type="checkbox"/>
<input type="checkbox"/>	Experiment 6	Joseph Mann	<input type="checkbox"/>
<input type="checkbox"/>	Experiment 8	Joseph Mann	<input type="checkbox"/>
<input type="checkbox"/>	Experiment 9	Joseph Mann	<input type="checkbox"/>
<input type="checkbox"/>	Experiment 10	Joseph Mann	<input type="checkbox"/>

Easier for model to understand data

- LLMs already understand normalized databases and explaining well-structured schemas takes less space in a context window.

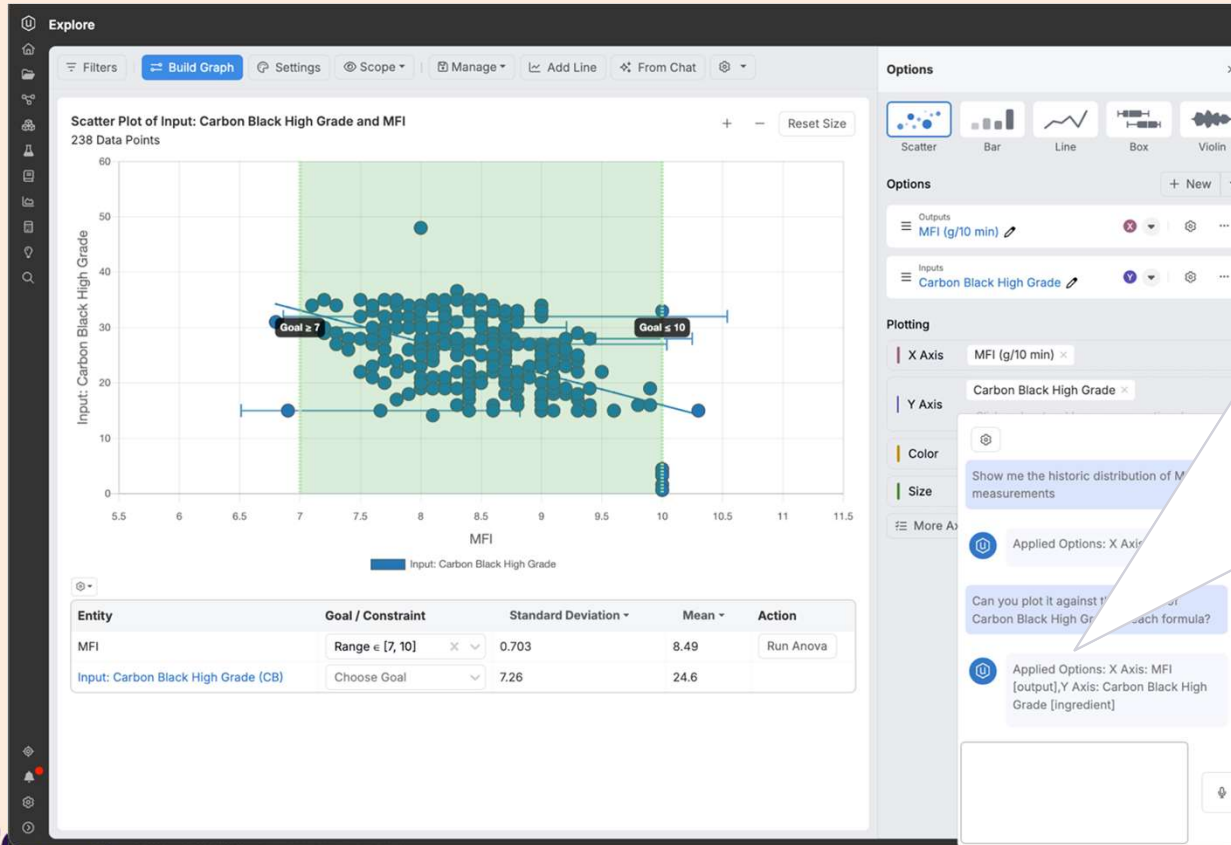
Interactions with The Structure

- LLMs understand the data structure and can insert and modify data according to the structure.

Output is Validated By The Data Model

- Interactions happen through controlled API with database and application-level validation.

Platform in Action



Settings
Close

Show me the historic distribution of MFI measurements

Applied Options: X Axis: MFI [output]

Can you plot it against the amount of Carbon Black High Grade in each formula?

Applied Options: X Axis: MFI [output], Y Axis: Carbon Black High Grade [ingredient]

Microphone

LLM: Ingredient Research

"I'm looking for alternatives to the ingredient Kronos 2220 to help with chalking performance."

KRONOS® 2220 is a rutile titanium dioxide (TiO₂) pigment produced via the chloride process. It is specifically designed for plastics applications requiring high outdoor durability and excellent optical properties. The pigment is surface-treated with aluminum, silicon, and polysiloxane compounds, enhancing its performance in various applications.

Key Properties:

- **High Durability:** Enhances the longevity of plastics exposed to outdoor conditions.
- **Brightness and Neutral Tone:** Provides very high brightness with a neutral undertone.
- **Tinting Strength and Opacity:** Offers high tinting strength and opacity, contributing to effective coverage.
- **Dispersibility:** Readily wetted and dispersed, facilitating ease of use in formulations.

	Name	Description	Chemical Info
+	Kronos 2190	A rutile titanium dioxide pigment with enhanced weather resistance and surface treatment, designed for outdoor applications with improved chalking resistance.	<chem>O=[Ti]=O</chem> Name TITANIUM DIOXIDE Molecular Weight 79.865 g/mol CAS Number 13463-67-7 SMILES O=[Ti]=O
+	Tronox CR-828	A chloride-process rutile titanium dioxide pigment with high durability and excellent resistance to chalking, suitable for exterior plastics and coatings.	<chem>O=[Ti]=O</chem> Name TITANIUM DIOXIDE Molecular Weight 79.865 g/mol CAS Number 13463-67-7 SMILES O=[Ti]=O

IRI 2026

**INNOVATORS
SUMMIT**



Data Summary Tools

- Summarize a set of data or project automatically to find key insights and make plots + tables

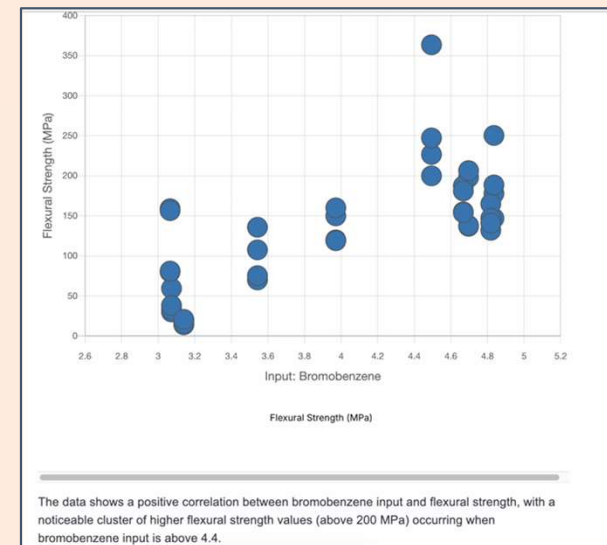
"I just started working on this project. Help me understand the data here "



Next let's analyze the trends in this data.

Insights on Experiments

1. There is a strong, positive correlation between New Polymer B (ingredient) and Flexural Strength (MPa) (output). Experiments with higher amounts of New Polymer B (e.g., 4.8–4.9) tend to show higher Flexural Strength values (e.g., 363.6, 226.7, 206.5 MPa), while lower or absent amounts correspond to lower Flexural Strength.
2. Higher levels of New Polymer C (ingredient) are associated with increased Elongation at Break (%) (output). For example, experiments with New Polymer C above 4.6 show Elongation at Break values of 59.96, 47.79, and 86.83%, which are among the highest in the dataset.
3. There is a negative correlation between UV Stabilizer 1 (ingredient) and Tensile Modulus (MPa) (output). Experiments with higher UV Stabilizer 1 (around 3.7–3.9) tend to have lower Tensile Modulus values (e.g., 1.49, 1.92, 1.66), while lower amounts (around 1.1) are present in experiments with higher Tensile Modulus (e.g., 19.45, 17.15 MPa).





Q&A

Thank You!

Questions?

Email: info@uncountable.com

Inquiries: www.uncountable.com/contact-us

www.uncountable.com

Want a Demo?



Scan The QR Code