Moneyball for Material$

AI for commercializing PFAS-free DWR

Vasav Sahni, PhD
Polymer scientist specializing in performance textile finishes and PFAS-free systems.

CONTENTS

Executive summary

PFAS-free DWR is a multi-variable optimization problem: finish chemistry, textile substrate variability, process windows, and durability must all work together.

AI is most valuable when it makes rigor scalable: it prioritizes the next best experiments, learns from failures, and reduces false positives that collapse during mill trials.

The highest-ROI AI pattern for coatings teams is small-data active learning (often Bayesian optimization) combined with disciplined testing metadata.

Industry proof points that exist: Bayesian optimization has been used to develop coating materials with fewer iterations; self-driving lab concepts demonstrate closed-loop active learning; and coatings companies have adopted materials informatics to accelerate reformulation and regulatory-driven ingredient swaps.

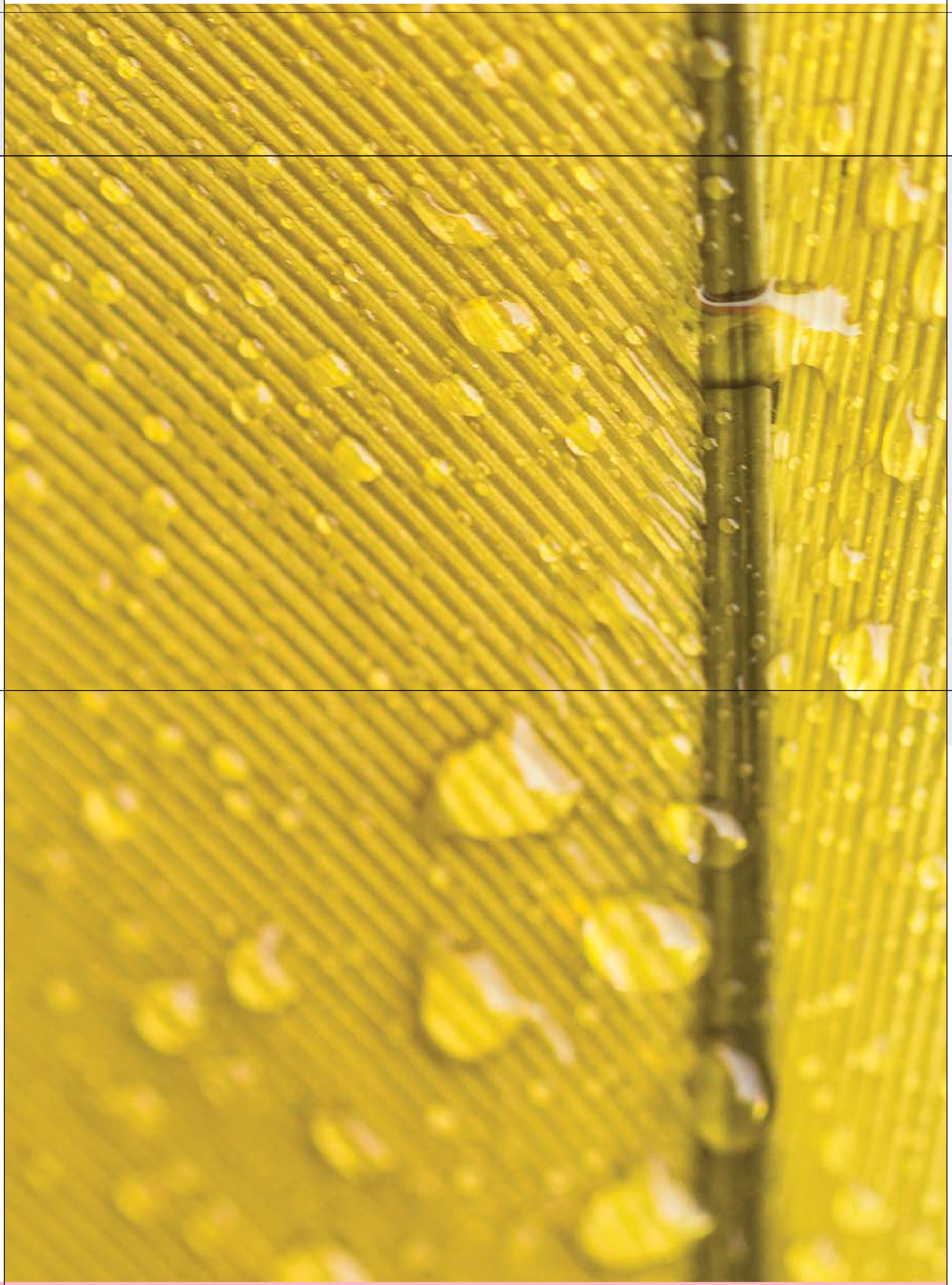
The 80/20 for PFAS-free DWR: standardize your test + process data, incorporate fabric readiness/contamination descriptors, then use active learning to steer formulation and process exploration toward robustness.

1.

Why AI now: the PFAS-free DWR equation is getting harder

In the first three papers in this series, we argued that PFAS-free DWR success depends on two realities: (1) rigor is the speed strategy, and (2) textile substrates are not homogeneous; contamination and variability are structural. AI becomes relevant not as a substitute for chemistry or testing, but to operationalize rigor at scale: deciding which experiments to run next, learning faster from noisy data, and reducing expensive mill-trial failures.

This matters because DWR performance is an interface problem: small changes in surface chemistry, textile preparation, and process conditions can materially change repellency and durability. Durable water repellent agents are commonly applied as emulsions and form thin surface modifications on fibers, preserving pore structure for breathability - an advantage that also increases sensitivity to residues and process drift.[1][2]



2.

What AI can (and cannot) do for PFAS-free DWR

AI adds the most value where the human workflow is weakest: searching large design spaces under uncertainty. Three practical value levers:

- **Prioritize experiments: recommend the next formulation/process trials that maximize learning per experiment (active learning / Bayesian optimization).**
- **Predict robustness: learn relationships between formulation + process + substrate metadata and outcomes such as spray rating and durability retention.**
- **Capture and reuse tacit knowledge: convert lab notes, test reports, and supplier specs into searchable structured data to avoid repeating mistakes.**

What AI will not do by itself: remove the need for standardized testing, clean metadata, and clear decision gates. A model trained on inconsistent labels or hero fabrics will confidently recommend the wrong next step.

Exhibit 1. AI options for coatings R&D and when they tend to pay off

AI OPTION	BEST FOR	TYPICAL DATA REALITY	WHAT IT CHANGES DAY-TO-DAY
Active learning / Bayesian optimization	Small-to-medium data formulation and process optimization	50-500 experiments	Suggests next experiments: reduces iterations
Supervised ML models	Predicting outcomes within a known space	200-2,000 experiments	Screens candidates: explains drivers of failures
Multi-objective optimization (Pareto)	Tradeoffs (durability vs hand feel vs breathability vs cost)	Clear objective definitions	Makes tradeoffs explicit: avoids single-metric false wins
Knowledge graph & LLM copilot	Institutional memory and faster troubleshooting	Documents and lab notes	Finds precedent; standardizes terminology
Self-driving labs/ closed-loop automation	High-throughput, standardized workflows	Automation & tight loops	Runs experiments continuously: scales learning

3.

The AI options landscape

Two patterns from the broader materials community are worth highlighting:

First, Bayesian optimization and related active learning methods are widely used to optimize expensive experiments with limited data.[3][4]

Second, the self-driving laboratory concept: automation plus autonomous experiment planning - has matured rapidly, with recent reviews describing its potential to accelerate materials discovery.[5][6]

4.

Evidence and proof points: what has already been done

A few grounded examples:

Bayesian optimization applied to coating development: published work describes Gaussian-process-assisted Bayesian optimization used to develop a zirconia-based coating material—use a surrogate model to guide trials rather than brute-force iteration.[3]

Closed-loop active learning in materials synthesis: Nature reported an autonomous laboratory where active learning proposes follow-up recipes when synthesis attempts fail.[7]

Coatings organizations adopting materials informatics: industry-facing materials informatics examples describe accelerating reformulation for regulatory compliance and shortening time-to-target by learning from historical formulation data.[8]

ML workflows for functional coatings: open-access research describes ML-assisted discovery and experimental validation in polymeric coatings.[9]

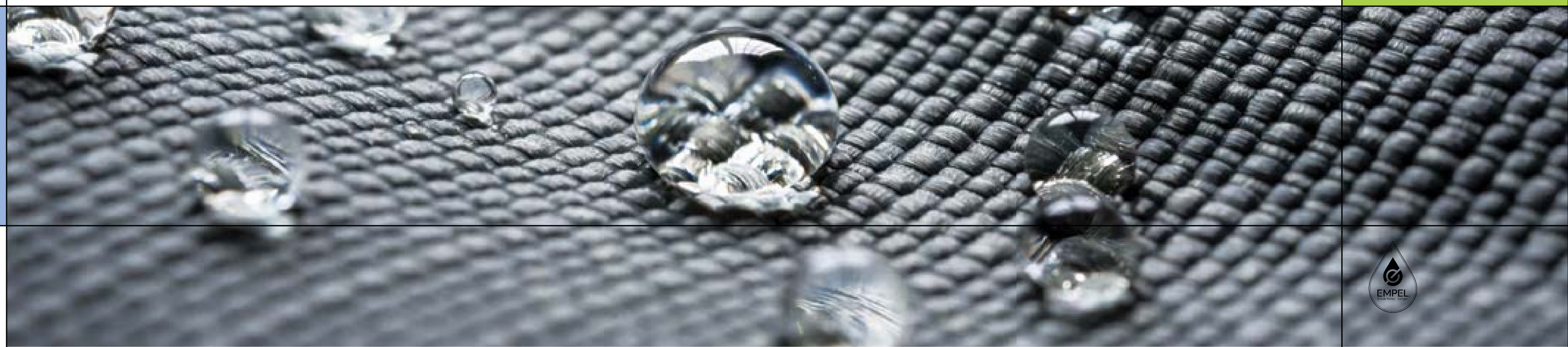
5.

The operating model: RIGOR + AI as a closed learning loop

In PFAS-free DWR, the objective is not to optimize a single lab metric; it is to maximize the probability of mill-scale success on variable fabrics. The most pragmatic approach is a closed learning loop compatible with small data and high uncertainty.

A five-step loop that most coatings teams can implement without a self-driving lab:

1. Define success and gates: map claims to end-use tiers (surface repellency vs penetration risk vs oil/soil exposure) and set pass/fail criteria.
2. Instrument the experiment: capture formulation, process conditions, fabric readiness, and contamination risk descriptors (not just outcomes).
3.S tart with DOE, then switch to active learning: DOE provides coverage; Bayesian optimization exploits what you learn next.[4]
4. Train for robustness: include representative fabric families and contamination challenge panels so the model learns real-world variability.
5. Make tradeoffs explicit: use multi-objective scoring to prevent single-metric false winners and to document why you chose a candidate.



6.

Minimum viable dataset (MVD): the fields that unlock value

Most teams already have data; it is typically missing the context that drives repeatability. The goal is an MVD that supports recommendations, troubleshooting, and scalability.

Minimum viable dataset (MVD): the fields that unlock value

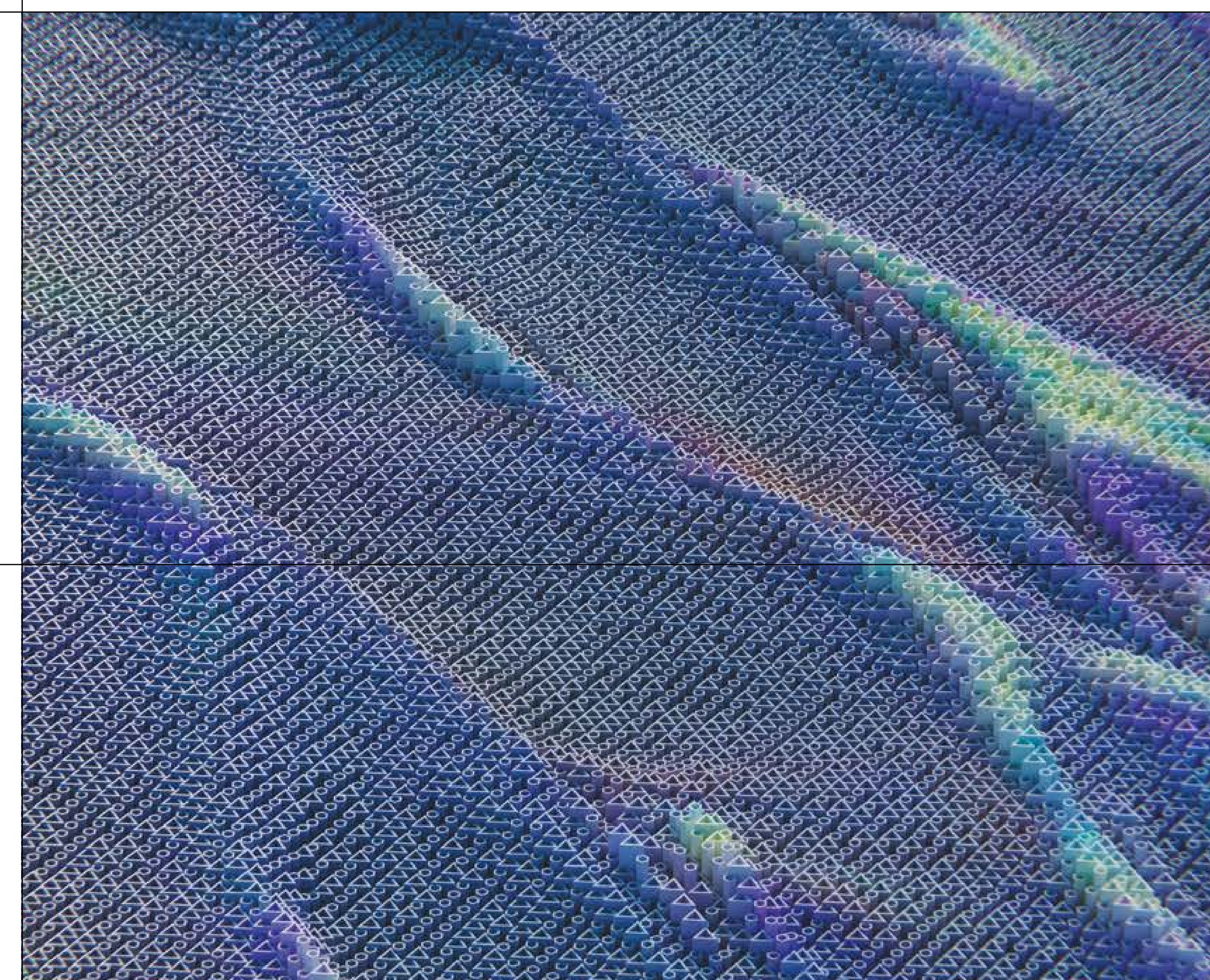
CATEGORY	RECORD AT MINIMUM	WHY IT MATTERS
Formulation	Ingredient ID's; solids; ratios	Enables structure-process-property learning
Process	Bath pH; pickup/add-on; dru/cure profile	Defines process widow and sensitivity
Textile substrate	Fiber/blend; construction; supplier/lot	Captures substrate-driven varience
Readiness / contamination	pH/alkalinity; extractables proxy; silicone exposure	Prevents mystery failures; links to contamination risk[2]
Outcomes	Spray rating; durability after wash; (oil repellency if needed)	Trains models on what customers experience

7.

Decision-making: picking the right AI approach

Match the sophistication of AI to the bottleneck and the data reality:

- If you have <100 experiments: prioritize standardization + DOE + searchable knowledge capture. Don't overfit models.
- If you have 100–500 experiments: deploy active learning / Bayesian optimization to recommend next trials and quantify uncertainty.[4]
- If you have >500 experiments across fabric families: add supervised models and a robustness score; treat domain shift as a first-class risk.
- If you can automate measurements and run high-throughput: consider closed-loop workflows; self-driving lab literature shows the trajectory.[5][7]



8.

The 80/20: where AI creates the biggest benefit in our space

The highest ROI actions are to:

- Enforce the MVD across labs and mills.
- Use active learning to optimize the next 10 experiments, not the next 1,000.
- Train for robustness by design: include substrate variability and contamination challenge panels so the model learns what breaks PFAS-free performance.

Active learning methods were built for this scenario: expensive experiments, limited data, and the need to balance exploration and exploitation, which are also emphasized in self-driving lab and materials optimization literature.[4][5]

9.

A 90-day implementation plan (minimum viable AI)

Most teams need a disciplined pilot that proves value.

Days 0–30: Define objectives and gates; implement the MVD; unify test methods; choose 2–3 representative fabrics including a worst-case substrate.

Days 31–60: Run an initial DOE; train a baseline model; start Bayesian optimization recommendations; review weekly with R&D + manufacturing.

Days 61–90: Add durability and contamination challenge panels; expand to a second fabric family; document process windows; define ready-for-mill-pilot criteria.

Selected references

[1] Holmquist, H. et al. Properties, performance and associated hazards of state-of-the-art durable water repellent (DWR) chemistry for textile finishing. Environment International (2016).
https://doi.org/10.1016/j.envint.2016.04.012

[2] Cotton Incorporated. ISP 1007: Water and Stain Repellent Finishing of Cotton Fabrics (PDF).
https://www.cottoninc.com/wp-content/uploads/2017/12/ISP-1007-Water-and-Stain-Repellent-Finishing-of-Cotton-Fabrics.pdf

[3] Park, S.M. et al. Gaussian process regression-assisted Bayesian optimization (zirconia-based coating development). (2023).
https://www.sciencedirect.com/science/article/pii/S2238785422020610

[4] Pai, S.M. et al. Machine learning applied to the design and optimization of chemical products and functional materials (includes Bayesian optimization). (2025).
https://www.sciencedirect.com/science/article/pii/S2949822824003472

[5] Tom, G. et al. Self-Driving Laboratories for Chemistry and Materials Science. Chemical Reviews (2024). Open access:
https://pmc.ncbi.nlm.nih.gov/articles/PMC11363023/

[6] Szymanski, N.J. et al. An autonomous laboratory for the accelerated synthesis of materials. Nature (2023).
https://doi.org/10.1038/s41586-023-06734-w

[7] Citrine Informatics. Materials Informatics for Coatings Formulations: Applied Machine Learning Strategies for Rapid Reformulation (PDF).
https://citrine.io/wp-content/uploads/2023/04/White-Paper-Materials-Informatics-for-Coatings-Formulations.pdf

[8] Liu, T. et al. Machine learning assisted discovery of high-efficiency self-healing epoxy coating for corrosion protection. npj Materials Degradation (2024) (PDF).
https://pure.tudelft.nl/ws/portalfiles/portal/175946281/s41529-024-00427-z.pdf

EMPEL is a PFAS-free DWR technology applied to the outer surface of textiles. Unlike traditional coatings that soak into the fabric and wash out quickly, EMPEL is pressed into the fibers – without using water or PFAS – providing long-lasting, high-performance water repellency.

Sheds water. Longer.

Contact info

4132 Jackie Rd
SE Rio Rancho , NM
87124

https: / /empel.g reen/#contact