

Human Oversight and Reproducibility in Autonomous Materials Research: Authority, Auditability and Failure Containment

Governance requirements for scientifically defensible autonomous experimentation

Juan Seguro

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Seguro. All Rights Reserved.

Correspondence: info@3Dresyns.com · www.3Dresyns.com

ER-410 · Engineering Review · Version 0.1 · 7 October 2026

Abstract

Autonomous materials research changes who or what selects experiments, but it does not remove scientific accountability. This review defines human oversight as an engineering control system with explicit authority boundaries, audit trails, intervention rules, independent replication and failure containment. Safety-focused self-driving-laboratory literature shows that increasing automation creates new requirements for safeguards and human-machine responsibility. For photopolymer development, the problem is amplified by coupled chemistry, light exposure, thermal behaviour, robotic handling and post-processing. The framework separates authorisation to propose, execute, modify and release; requires every autonomous decision to remain reconstructable; and defines conditions under which the system must stop rather than continue learning. Reproducibility is treated as a property of the entire closed loop, not merely repeated execution of the same software.

Keywords: human oversight · autonomous materials research · reproducibility · audit trail · authority · failure containment · self-driving laboratory · photopolymer

1. Autonomy does not transfer accountability

Self-driving laboratories automate both experimental execution and experiment selection [1,2]. Safety-focused analyses emphasise that autonomous systems require explicit safeguards and governance rather than an assumption that automation is inherently safer [3]. The scientific author remains responsible for the claims drawn from the resulting evidence. Human oversight should therefore be designed into the workflow. A person who only sees the final optimum is not meaningfully supervising the experiment-selection process.

2. Authority matrix

Four authorities should be distinguished: propose an experiment, execute an experiment, modify the search space or method, and approve a scientific or release claim. These authorities need not belong to the same agent or person. An autonomous controller may be authorised to propose and execute within a frozen envelope while changes to hazardous materials, hardware limits, measurement methods or publication claims remain human-controlled.

3. Auditability

Every autonomous action should be reconstructable from immutable or version-controlled records: input data snapshot, model and software version, constraints, proposed candidate, selection rationale, execution record, raw measurements, preprocessing and resulting update. Auditability is necessary for scientific reproducibility because an adaptive campaign cannot be reproduced from the final recipe alone. The sequence of observations influences later choices.

Table 1. Authority matrix

Action	Automation may act when	Human authority remains required for
Propose experiment	inside frozen domain	domain expansion
Execute experiment	all interlocks satisfied	override of safety limit
Modify method/model	version-controlled test environment	production scientific method
Approve claim	never autonomously final	scientific/publication/release claim

4. Reproducibility of adaptive campaigns

Repeating an adaptive campaign does not necessarily generate the same sequence because measurement noise and probabilistic selection alter the path. Reproducibility should therefore be defined at multiple levels: executor repeatability, measurement repeatability, algorithmic replay from a frozen dataset, and convergence of scientific conclusions across independent campaigns. Random seeds and software environments should be recorded where stochastic algorithms are used, but identical random seeds do not substitute for independent experimental replication. Metrological studies of liquid-handling robotics show why automated execution itself must remain reproducible and auditable [4].

5. Human intervention

Interventions are not failures if they are recorded. A chemist may recognise phase separation, contamination or a physically implausible proposal that automated sensors missed. The intervention should be logged with reason, scope and consequence for the dataset. Unrecorded manual rescue is scientifically damaging because it creates hidden selection. The published campaign can then appear more autonomous and reproducible than it actually was.

6. Failure containment

Failure containment includes software validation, hardware interlocks, spill or temperature protection, safe shutdown, candidate-domain limits and quarantine of anomalous data. A learning system should not convert a hardware fault into a new scientific relationship. After a containment event, restart criteria should establish that equipment and measurement references are back in control before new data re-enter the model.

Table 2. Reproducibility layers

Layer	Question	Evidence
Executor	Is the same command delivered consistently?	repeatability/control sample
Measurement	Does the method reproduce the measurand?	replicates and uncertainty
Algorithmic replay	Can the decision sequence be reconstructed?	versioned data/code/logs
Scientific replication	Does the conclusion recur independently?	new preparation/campaign

7. Change control

Changes to code, model architecture, acquisition function, calibration, hardware, raw-material lot policy or measurement method can alter the scientific system. They should be versioned and assessed for whether old and new observations remain poolable. A campaign spanning a material change without a recorded change point risks treating two different experimental populations as one.

8. Independent replication

Key discoveries should be independently remade and measured outside the adaptive sequence. Where practical, a second operator, preparation route or instrument provides stronger evidence that the result is not tied to one automation path. Replication should target the claimed conclusion, not merely a single numeric response. If the claim is a trade-off frontier, several representative points may need confirmation.

9. Publication and evidence transparency

A publication should disclose the bounds of automation, human interventions, failed experiments, exclusions, software/model versions and independent confirmation. It should not present the controller as an author of scientific judgement; responsibility for interpretation remains with the human author. The absence of hidden assistance or unrecorded data curation is more important than the degree of autonomy itself. A partially automated but fully auditable study can be scientifically stronger than a nominally autonomous black box.

10. Qualification decision boundary

Human oversight and autonomous experimentation are complementary. Oversight defines authority, detects conditions outside machine competence, protects safety and preserves accountability. Autonomous execution can then improve speed and consistency within that bounded space. For photopolymer materials research, the qualification target is a reproducible human-governed closed loop whose failures, interventions and evidence lineage are visible enough to support independent scrutiny.

11. Competence and escalation design

Oversight depends on competent review, not merely the presence of a person. The governance plan should define which decisions require chemistry, metrology, process, safety or statistical expertise and how uncertain cases are escalated. A generic approval click does not constitute meaningful scientific review. Escalation criteria should be objective where possible: out-of-domain score, failed reference control, safety-interlock event, disagreement among measurements, model update beyond a declared threshold, or proposal outside the validated chemistry family.

12. Incident and near-miss learning

Near misses are valuable evidence in autonomous laboratories because they reveal weak containment before a serious event occurs. The system should preserve near-miss records and link them to subsequent changes in constraints, hardware or procedures. Repeated near misses of the same class indicate that the control hierarchy is not addressing the underlying cause. Scientific data from an incident should be quarantined until it is established that the experiment remained within measurement validity. A safe shutdown does not guarantee that the partial measurement is suitable for model training.

13. Reproducibility across laboratories

Cross-laboratory reproducibility requires more than sharing software. Executor calibration, raw-material identity, environmental conditions, process hardware and measurement methods all affect the closed loop. The reproducible object is therefore the controlled experimental specification plus its tolerances, not a code repository alone. Independent laboratories may follow different adaptive paths and still support the same scientific conclusion. Evaluation should compare final causal or performance claims with uncertainty, while analysing why candidate sequences diverged.

14. Evidence gaps and research priorities

Autonomous materials research needs common reporting standards for authority, interventions, containment events, software changes and independent confirmation. Such reporting would allow readers to distinguish genuine autonomy from workflows that depend on substantial undocumented manual judgement. For photopolymer research, challenge studies should deliberately inject executor faults, drifting references, unsafe proposals and conflicting measurements to test whether governance mechanisms detect, contain and document the problem before it contaminates the scientific record.

References

1. Tom G, Schmid SP, Baird SG, Cao Y, Darvish K, Hao H, Lo S, Pablo-García S, Rajaonson EM, Skreta M, Yoshikawa N, Corapi S, Akkoc GD, Strieth-Kalthoff F, Seifrid M, Aspuru-Guzik A. Self-Driving Laboratories for Chemistry and Materials Science. *Chemical Reviews*. 2024;124(16):9633-9732. doi:10.1021/acs.chemrev.4c00055.
2. Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis*. 2023;2:483-492. doi:10.1038/s44160-022-00231-0.
3. Leong SX, Griesbach CE, Zhang R, Darvish K, Zhao Y, Mandal A, Zou Y, Hao H, Bernales V, et al. Steering towards safe self-driving laboratories. *Nature Reviews Chemistry*. 2025;9:707-722. doi:10.1038/s41570-025-00747-x.
4. Hambidge T, Corless S, Cowen S, Short M, Hopley C, Sears P. Pitfalls and Inherent Biases in Liquid Handling Robotics: Investigations in Automation for SI Traceable Measurements. *Analytical Chemistry*. 2026;98(22):16107-16116. doi:10.1021/acs.analchem.6c00078.

Competing interests

The author is Founder, CEO and CTO of 3Dresyns by Resyner Technologies S.L.; ER-410 contains no comparative recommendation between commercial products.

Data availability

ER-410 is a literature-based engineering review; no new experimental dataset was generated.