

Multimodal Foundation Models for Photopolymer Engineering: Text, Spectra, Images, Process Data and Evidence Grounding

A provenance-preserving framework for combining heterogeneous engineering evidence

Juan Segurola

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Segurola. All Rights Reserved.

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

ER-404 · Engineering Review · Version 0.1 · 7 October 2026

Abstract

Photopolymer engineering generates heterogeneous information: chemical and process text, spectra, rheology curves, optical images, build logs, dimensional measurements and service data. Multimodal foundation models offer mechanisms for joint representation, retrieval and cross-modal association, but current materials-science evidence is not equivalent to a validated photopolymer decision system. This review defines a qualification architecture in which each modality retains identity, units, provenance and measurement uncertainty. The model may assist retrieval, alignment or prediction, but source evidence remains the controlling record. The review distinguishes representation learning from evidence grounding, addresses domain shift and out-of-distribution inputs, and proposes claim-level tests for models used in formulation or process decisions. The central requirement is that a multimodal output must be traceable back to the measured or cited evidence that supports it.

Keywords: multimodal foundation model · photopolymer engineering · spectra · imaging · process data · provenance · evidence grounding · domain shift

1. Scope and evidence boundary

Materials-science foundation models have begun to combine text, structure and images, and recent work has evaluated cross-modal image matching in materials microscopy [1,2]. These studies establish technical feasibility for multimodal representation, but they do not establish a validated foundation model for photopolymer formulation or qualification. Any photopolymer implementation therefore begins as a transfer problem. The relevant engineering question is not whether heterogeneous data can be embedded together. It is whether combining them preserves the measurement context needed to make a defensible decision.

2. Modalities in a photopolymer record

Text can contain formulation metadata, procedures, deviations and literature claims. Spectra can encode absorption, conversion or chemical state. Images can represent cure geometry, defects, surface condition or microstructure. Process logs describe irradiance, exposure sequence, temperature and machine events. Mechanical, optical, electrical and dimensional results add structured numeric outputs. Each modality has its own sampling geometry, resolution, units and uncertainty. A joint model should not erase those differences. The source modality and transformation history should remain recoverable for every derived feature.

3. Alignment before fusion

Multimodal fusion assumes that records correspond to the same specimen, region, time or state. Misregistration can create false relationships even when each measurement is individually valid. The microscopy literature demonstrates that image matching across modalities and length scales remains difficult in low-overlap cases [2]. For photopolymer studies, alignment can be temporal as well as spatial: a spectrum before cure, a thermal trace during cure and a mechanical measurement after post-cure refer to different states. Their linkage must be explicit rather than implied by a shared sample name.

Table 1. Modality control requirements

Modality	Critical metadata	Typical failure
Text	source; version; section; date	claim detached from source
Spectrum	instrument; baseline; path; units; state	false chemical similarity
Image	scale; region; illumination; specimen	misregistration or scale leakage

4. Evidence grounding

A model-generated summary or prediction is not itself primary evidence. The controlling evidence remains the source spectrum, image, log, measurement record or cited publication. Grounded systems should therefore return source pointers, specimen identifiers and relevant data windows alongside any conclusion. When literature text is included, bibliographic provenance should be retained at passage level. A model should not combine claims from different material families into a single statement without indicating the source boundaries.

5. Representation versus interpretation

A shared embedding can support retrieval or similarity, but geometric proximity in embedding space does not establish chemical equivalence. Two spectra may look similar while corresponding to different path lengths, concentrations or baselines. Two defect images may look similar while arising from different failure mechanisms. Interpretation should therefore be layered: representation for indexing; task-specific calibrated models for prediction; and separate evidence rules for engineering claims.

6. Domain shift and out-of-distribution data

Foundation models trained on broad data can fail when confronted with specialized photopolymer images, spectra or terminology. Recent materials work explicitly evaluates out-of-domain generalisation rather than assuming it [1,2]. Photopolymer deployment should do the same. OOD detection should be treated as a gate. When a modality is outside the validated domain, the system should defer, request additional measurement or route the record for expert review rather than force a confident prediction.

Modality	Critical metadata	Typical failure
Process log	clock; machine; calibration; event schema	wrong temporal alignment
Property data	method; specimen geometry; uncertainty	incomparable endpoints

Table 2. Qualification tests for multimodal models

Test	Purpose	Release criterion concept
Domain holdout	measure transfer	predefined performance per domain
OOD challenge	test deferral	reject unsupported inputs
Missing-modality test	robustness	no silent fabrication
Provenance audit	evidence grounding	source retrievable for claim
Leakage audit	independence	no specimen/batch duplication across splits

7. Missing and conflicting modalities

Real engineering records are incomplete. A missing spectrum should not be silently imputed as if measured. Conflicting evidence is also meaningful: a nominally high conversion can coexist with poor dimensional fidelity or residual extractables because the endpoints probe different states and mechanisms (see ER-027, ER-033 and ER-035). Multimodal systems should expose missingness and conflict. The engineering decision layer can then define whether a missing modality is tolerable or whether it blocks the claim.

8. Evaluation architecture

Model evaluation should use specimen-level splits that prevent near-duplicate images or spectra from leaking between training and test sets. Where batches, printers or material families are relevant, leave-one-domain-out tests are more informative than random splits. Evaluation should include retrieval accuracy, calibration, OOD rejection, robustness to missing modalities and traceability of cited evidence. A high aggregate accuracy can hide unacceptable failures in a safety-critical subgroup. Uncertainty-aware bridging between simulation and experiment provides a relevant adjacent evaluation precedent [3].

9. Governance and records

Every model output used in an engineering decision should be bound to model version, input record hashes or stable identifiers, preprocessing version and confidence or uncertainty information. If the model is updated, prior decisions remain linked to the version that produced them. Human review is required where the system proposes a claim that is not directly supported by retrievable evidence or where modalities disagree on a release-relevant endpoint.

10. Qualification decision boundary

Multimodal foundation models can improve navigation and integration of photopolymer evidence, especially where measurements are heterogeneous. Their value depends on preserving provenance and state rather than compressing everything into an opaque representation. The appropriate qualification target is a grounded, auditable decision aid. It should make evidence easier to interrogate, not replace the evidence record.

11. Spectral and image preprocessing as part of the model

Baseline correction, normalization, resampling, denoising, segmentation and image resizing can materially alter the information presented to a multimodal model. These operations are not neutral data cleaning. They define which physical variation is retained and can introduce leakage if parameters are fitted using the full dataset before the train-test split. Preprocessing must therefore be versioned and validated per modality. Raw data should remain available so that a surprising model output can be traced back through each transformation. Where multiple preprocessing pipelines are plausible, robustness to those choices should be tested.

12. Retrieval and citation integrity

For literature-grounded use, retrieval quality should be evaluated independently from generation or synthesis quality. A system can produce a fluent answer from the wrong paper. Each extracted claim should carry source identity, location and version, and the evaluation should include adversarially similar documents from neighboring material families to test whether the model preserves scope boundaries. Bibliographic metadata should be treated as structured data rather than reconstructed from model memory. DOI, title, authors, journal and year should come from a controlled record or verified source. Missing metadata should remain missing until verified rather than being completed probabilistically.

13. Cross-modal conflict as scientific information

When modalities disagree, the conflict may indicate a real mechanism. For example, spectroscopy can indicate high average conversion while imaging reveals geometric defects; thermal data can indicate a transient hot spot while final dimensions remain acceptable; rheology can appear stable while sedimentation changes local filler concentration (see ER-027, ER-351, ER-356, ER-357 and ER-359). Fusion should not force these signals into consensus. A useful system should surface conflicts and suggest the additional measurement that would discriminate among explanations. This turns multimodal integration into experimental reasoning rather than majority voting among sensors.

14. Evidence gaps and research priorities

Photopolymer-specific multimodal benchmarks should link raw spectra, calibrated optical fields, images, formulation metadata, process logs and independently measured properties at specimen level. Public datasets need explicit split rules that prevent batch and specimen leakage, and should include missing modalities and negative results. Foundation-model performance should be compared with narrower task-specific models. A larger general model is not automatically the engineering optimum if a smaller calibrated model provides better provenance, uncertainty and domain control for a specific release decision.

References

1. Tang Y, Xu W, Cao J, Gao W, Farrell S, Erichson B, Mahoney MW, Nonaka A, et al. A multimodal large language model for materials science. *Nature Machine Intelligence*. 2026;8:588-601. doi:10.1038/s42256-026-01214-y.
2. Durmaz AR, Lamb JD, Echlin MP, Pollock TM. Foundation models for multimodal image data fusion in materials science. *Frontiers in Materials*. 2026;13:1815017. doi:10.3389/fmats.2026.1815017.
3. Chen J, Long T, Wall M, Hufnagel T, Chen W. Uncertainty aware machine learning for bridging simulation and experiment in high throughput materials characterization. *Scientific Reports*. 2026;16:20837. doi:10.1038/s41598-026-51212-8.

Competing interests

The author is Founder, CEO and CTO of 3Dresyns by Resyner Technologies S.L.; ER-404 contains no comparative recommendation between commercial products.

Data availability

ER-404 is a literature-based engineering review; no new experimental dataset was generated.