

Software Supply Chains Are Dead: Use-Case-Oriented Regeneration

Tanmay Singla ✉ 

Purdue University, West Lafayette, IN, USA

James C. Davis ✉ 

Purdue University, West Lafayette, IN, USA

Abstract

Modern software development relies on an increasingly doubtful premise: that the up-front implementation savings from adopting a dependency outweighs the maintenance costs. Two changes are reshaping the build-*vs.*-reuse calculus: software supply chain attacks have raised the cost of external reliance, while generative AI has lowered the cost of local implementation. In this vision paper, we discuss *use-case-oriented regeneration* as a new software sourcing paradigm that shifts the supply chain from external trust to local verification. We evaluate an agentic workflow that synthesizes only the specific slice of dependency functionality that a repository exercises. Our measurements across 180 repository-dependency pairs suggest that this approach is feasible: the replacements preserve 99.8% of repository-observed behavior across baseline validation checks and reduce the exported API surface by 93%. Software sourcing may evolve toward verifiable repository-specific code synthesis, especially when the required functionality is narrow, stable, and well tested.

2012 ACM Subject Classification Software and its engineering → Software libraries and repositories; Software and its engineering → Software verification and validation; Computing methodologies → Artificial intelligence

Keywords and phrases Software supply chain, dependency reuse, code generation, AI coding agents

Digital Object Identifier 10.4230/LIPICs.ESEM.2026.64

Category Emerging Results, Vision & Reflection Track Paper

Supplementary Material

Dataset (Replication Package): <https://doi.org/10.5281/zenodo.20370439> [29]

Funding We acknowledge support from the US National Science Foundation under award #2541917.

1 Introduction

For over two decades, software reuse has been increasingly organized around package-level dependency adoption. When software engineers need functionality, the default posture is to depend on upstream packages, even when the repository uses only a small part of that package’s behavior. This artifact-centric model helped make reuse inexpensive and scalable [39, 41], but it also creates a persistent granularity mismatch: downstream repositories adopt, validate, and maintain entire upstream artifacts to obtain narrow pieces of functionality. That mismatch has become increasingly costly as repositories accumulate transitive dependency graphs and inherit upstream vulnerabilities, maintenance decisions, compatibility constraints, and ecosystem-level risks [2, 21, 13]. At the same time, generative AI and coding agents are lowering implementation costs as they become more capable of inspecting repositories, executing validation tools, and iteratively refining code [40, 36, 11, 4]. Together, these changes reopen the build-*vs.*-reuse decision: for narrow, stable uses of a dependency, relying on the full upstream artifact may no longer be the best choice.

We envision *use-case-oriented regeneration* as a new software sourcing paradigm that shifts dependency management from external trust to local verification. Instead of taking on a full dependency, an engineer can now use an agentic workflow to synthesize only the



© Tanmay Singla and James C. Davis;
licensed under Creative Commons License CC-BY 4.0

20th International Symposium on Empirical Software Engineering and Measurement (ESEM 2026).

Editors: Robert Feldt, Maria Paasivaara, Daniel Mendez, Stefan Wagner, and Marvin Muñoz Barón;

Article No. 64; pp. 64:1–64:14



Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

specific functionality that is actually needed as a local, repository-owned implementation. Building on prior regeneration work, we shift the unit of analysis from the upstream package to the *repository-dependency pair*. We report on the feasibility of this regenerative paradigm, evaluating regeneration across 180 repository-dependency pairs spanning nine diverse JavaScript/TypeScript dependencies. Our results show that agentic workflows can successfully replace full dependencies with narrow, local implementations while preserving 99.8% of repository-observed behavior across baseline validation checks. When successful, these generated replacements shed the unused code, transitive dependencies, and maintenance obligations that accompany artifact-centric reuse.

The shift to regeneration creates new responsibilities – it moves the trust problem from evaluating a persistent upstream maintainer [12] to validating the correctness and maintainability of locally synthesized code. This paradigm shift introduces open challenges regarding validation coverage, semantic equivalence, and how developers can establish trust in regenerated code. In sum, this paper makes the following contributions:

- We propose *use-case-oriented regeneration* as a repository-centric sourcing strategy that competes with conventional full-artifact reuse. We demonstrate its feasibility through an evaluation of 180 repository-dependency pairs, observing behavioral preservation with substantial footprint reduction.
- We outline a research agenda for synthesize-and-verify reuse, focusing on appropriateness, behavioral validation, provenance, and refresh practices for generated replacements.

2 Background and Motivation

2.1 Dependency-Mediated Reuse and the Build-vs.-Reuse Tradeoff

Package ecosystems have made reuse cheap by shifting functionality into upstream artifacts. Instead of building every component from scratch, developers assemble applications from packages distributed through ecosystems such as npm, PyPI, Maven, and Cargo [39, 41]. Industry reports estimate that open-source components underpin most modern software – Black Duck reports that 96% of audited commercial codebases contain open-source components [16, 33], and the Linux Foundation suggests a substantial footprint, with OSS comprising 70–90% of modern software. Third-party package dependencies are a primary mechanism by which software functionality is obtained, shared, and maintained.

Package reuse imports obligations beyond the functionality used. A downstream repository inherits not only the immediate functionality of a dependency, but also its transitive dependency graph, release practices, maintenance state, compatibility constraints, and vulnerabilities [20, 34, 10]. Dependencies evolve independently of the downstream projects that rely on them. A package update may fix a vulnerability, introduce a breaking change, or require downstream adaptation, while failing to update may leave a project exposed to known defects [7, 9]. Package ecosystems have also become attack surfaces: Sonatype reports more than 1.23 million known malicious open-source packages across major registries, including more than 454,000 newly identified during 2025 [32]. Dependency reuse is therefore a sourcing, maintenance, and assurance commitment.

The insight underlying this paper is that *dependencies are often adopted at a coarser granularity than they are used*. Package ecosystems distribute reusable artifacts as general-purpose packages, but a downstream repository may exercise only a small subset of a package’s API or behavior. The repository must still adopt, validate, and maintain the full artifact to obtain that subset. This mismatch motivates us to move beyond reactive reduction techniques and reconsider whether continued runtime reliance on full-package artifacts remains necessary for narrow dependency use cases [38, 1, 17].

This granularity mismatch changes the build-*vs.*-reuse tradeoff. Software sourcing is traditionally framed as a choice between implementing functionality locally or obtaining it from an external component [3]. Building locally gives the downstream project control over implementation, integration, review, and evolution, but requires design, implementation, and testing effort. Reusing an external package reduces the initial cost of acquiring functionality, but introduces costs related to integration, assurance, updates, and dependency management [15]. The tradeoff is therefore not only whether code exists, but who owns the implementation, how it is validated, and how it evolves.

The software engineering community is biased toward reuse; even the US Department of Defense approves of open-source software. We believe generative systems demand that we revisit that tradeoff. Large language models can inspect repositories, execute validation tools, and iteratively refine code, reducing the acquisition cost of local implementation by generating code for the surrounding repository context [36, 40, 19]. Reuse becomes less attractive when one’s repository uses a narrow and stable slice of the package, or when the package introduces transitive dependencies and security exposure that exceed the value of the functionality it provides [25]. By lowering the cost of code creation, generative systems expose another sourcing option: regenerating the verified slice of dependency functionality that the repository needs.

2.2 Motivating Example: A Case Study in Generative Sourcing

For a concrete illustration of this concept, let us share a recent experience. We were developing an automated accessibility tool [5] that needed to manipulate mathematics in Microsoft Office documents. The system had MathML and needed to render it as OMML for native support in Microsoft Office. Under the traditional sourcing model, we initially treated this as a dependency-selection problem. MathML-to-OMML conversion is not a trivial formatting task: it requires handling structured mathematical constructs, including operators, fractions, matrices, fences, scripts, annotations, and other element types. We therefore assessed existing implementations, including open-source converters, document-editor subprocesses, and commercial options.

Some candidates were plausible under the conventional reuse calculus: credible provenance, acceptable licensing, or good pilot behavior on our fixtures. However, that calculus assumes that the costs of taking on the dependency are lower than the cost of developing it ourselves. In our case, adding an external converter would have entailed more elaborate component vetting, licensing concerns, and bloating our Google Cloud images with the JVM.¹ We therefore considered an alternative: using an AI coding agent to implement the converter locally. We asked Claude about the complexity of MathML, which it characterized as involving roughly thirty relevant node types; and about the feasibility of a straightforward tree-transducer approach (whether each type could convert cleanly to OMML without reasoning over an intermediate representation). Receiving positive answers to each, we asked Claude to construct a MathML-to-OMML library of our own, backed by fixtures and validation against Office’s OpenXML SDK. After Claude spent an hour on it, the resulting library worked on the first try with our validation fixtures. Claude’s library has 1.4 KLOC, accompanied by 1.5 KLOC of unit and property-based tests. On inspection, the library code is of good quality – structured and concise. The library has been in use for the past two months with no reported functional defects. It was refactored once to extract helpers for sharing, and touched five times by repo-wide automated code modifications.

¹ Image bloat would reduce our service’s peak scalability under our current Cloud tenant memory quota.

■ **Table 1** Comparison of dependency reduction and replacement strategies. Use-case-oriented regeneration differs because it targets the repository-used slice of functionality as the unit of sourcing and replacement.

Approach	Starts from	Main goal	Relies on full upstream artifact?
Debloating	Adopted artifact	Remove unnecessary code and reduce attack surface	Yes, typically
Tree shaking / DCE	Adopted artifact	Remove unreachable shipped code	Yes
Vendoring	and build graph Upstream artifact	Internalize code for local control and patching	Yes, conceptually
Full regeneration (<i>e.g.</i> , HARP, Lexo)	Upstream artifact	Synthesize a full replacement of the package	No
<i>Use-case-oriented regeneration</i>	Repository– dependency pair	Replace repository-used functionality with local impl.	Not necessarily

This anecdote is not evidence that every converter should be regenerated. Rather, it illustrates the changed decision problem for package selection when the needed behavior is bounded and testable. *Is relying on existing packages still preferable once agent-assisted local synthesis makes a small, auditable, product-specific implementation cheap enough to create?*

2.3 Dependency Reduction and Internalization Strategies

Several existing strategies already attempt to mitigate the cost or exposure associated with overly broad dependencies. Debloating, for instance, removes unnecessary code from software artifacts or deployed systems while attempting to preserve required behavior [26]. Tree shaking and dead-code elimination remove code that is statically unreachable from a particular build, bundle, or execution path [37, 18]. Vendoring internalizes a dependency by copying the upstream source code directly into the downstream repository, often to improve control, reproducibility, or availability [24]. These strategies are highly useful because they explicitly recognize that the full adopted dependency often exceeds what a downstream project actually needs.

More recently, *full regeneration* strategies, such as HARP [35] and Lexo [14], have emerged to eliminate supply-chain attacks by leveraging program inference and LLMs to synthesize complete, semantically equivalent replacements for upstream packages.

However, as highlighted in Table 1, these strategies generally begin *after* a dependency has already been adopted as an upstream artifact, or they treat the *upstream package* as the fundamental unit of analysis. Debloating and tree shaking reduce unused code from an existing dependency, while vendoring moves upstream source into the downstream repository but still relies on the upstream implementation as the starting point. Full regeneration attempts to replace the artifact entirely, but still aims to synthesize a general-purpose, 1-to-1 package equivalent of the third-party library.

3 Use-Case-Oriented Regeneration

We propose *use-case-oriented regeneration* to address the structural liabilities of artifact-centric reuse. We define this approach as the local synthesis of only the specific dependency functionality that a downstream repository actually exercises, treating the *repository–dependency pair* as the fundamental unit of replacement. Crucially, this targets *repository–*

observed equivalence rather than full semantic equivalence with the upstream package. A generated replacement implements only the behaviors directly exercised by the repository, allowing unused API surfaces to be discarded, and success to be evaluated by whether the repository’s validation artifacts continue to pass.

In this context, we define a “use case” as the specific, bounded slice of a dependency’s API that is actually invoked by the downstream repository’s first-party code. This functionality is strictly bounded by the execution paths, configurations, and arguments exercised by the repository’s existing validation artifacts (e.g., its test suite). Therefore, the goal is not to broadly recreate the dependency, but to synthesize a minimal local implementation that perfectly satisfies these observed boundaries.

We can view use-case-oriented regeneration as a projection problem. Let P denote an upstream package, and let R denote a downstream repository that uses P . The repository induces an observation context O_R : the set of calls, inputs, configurations, validation artifacts, and execution paths through which R exercises P . Unlike full regeneration [35, 14], which seeks a replacement P' such that $P' \equiv P$ over the full package semantics, *use-case-oriented regeneration* seeks a local implementation $L_{R,P}$ such that

$$\pi_{O_R}(L_{R,P}) \equiv \pi_{O_R}(P),$$

where the projection π_{O_R} projects package behavior onto the behaviors observable through R ’s use of the dependency. Thus, the relevant equivalence class is defined by the repository–dependency pair, not by the upstream artifact alone.

This paradigm changes the locus of supply chain trust. In the current model, developers rely on external signals – such as package identity, ecosystem provenance, and maintainer reputation – that are often decoupled from the actual code executed [8, 6, 30]. Regeneration reduces runtime reliance on these external signals, shifting the trust problem from evaluating an upstream ecosystem to the *local verification* of synthesized code. The repository owner must therefore establish adequate evidence for the correctness, test coverage, and maintainability of a hyper-specialized local implementation.

4 Study Design

This section describes the empirical study used to evaluate use-case-oriented regeneration. We evaluate whether an agentic workflow can regenerate the functionality actually used by a client repository across a deliberately varied benchmark of repository–dependency pairs. Our study is a proof-of-concept: each successful regeneration provides evidence that a bounded slice of dependency functionality can be replaced by validated local code, while failures expose limits of the current workflow. Because regeneration is repository-specific and judgment-laden (cf. Section 2.2), the fully automated workflow described here is illustrative rather than prescriptive.

4.1 Research Questions and Evaluation Metrics

We ask three research questions, each paired with an evaluation metric:

RQ1 (Feasibility). Given a client repository and a dependency it uses, can an agent regenerate the exercised dependency functionality to remove first-party runtime reliance?

Rationale: This evaluates whether generative AI can practically alter the traditional build-vs.-reuse tradeoff.

Metric: Successful removal of the target dependency from runtime code and the successful integration of an executable local replacement.

RQ2 (Behavioral Preservation). To what extent do generated implementations preserve repository behavior across repository–dependency pairs?

Rationale: Because regeneration shifts the trust model from upstream reputation to local verification, this evaluates how well generated code satisfies existing repository assumptions.

Metric: The aggregate pass rate of the repository’s baseline validation checks (tests, builds, linters) before and after regeneration.

RQ3 (Size and Footprint). How do generated implementations compare with the original dependency in terms of size and dependency footprint?

Rationale: This contrasts use-case-oriented regeneration with reduction techniques like debloating, establishing whether repository-specific synthesis meaningfully reduces API exposure.

Metric: The measured reduction in exported API surface and implementation footprint compared to the original upstream package.

Taken together, these metrics evaluate both whether regeneration works as a repository transformation and what kind of artifact it produces when it succeeds. This distinction is critical: a generated replacement may pass repository validation while still raising concerns about untested behavior or maintainability.

4.2 Subject Dependencies and Repositories

We evaluate use-case-oriented regeneration on nine JavaScript/TypeScript dependencies selected to vary in size, role, popularity, and expected regeneration difficulty. The dependencies span small utility packages, medium-sized libraries, and larger packages with broader API surfaces. This selection is not intended to be statistically representative of all npm packages; instead, it is designed to exercise regeneration across a range of dependency types for which repository-specific replacement may plausibly range from straightforward to difficult [39, 41].

■ **Table 2** Target dependencies used in the study, grouped by approximate package size.

Size tier	Dependency	Type / role	Pairs
Small	nanoid	Identifier generation	20
Small	change-case	String case conversion	20
Small	chalk	Terminal string styling	20
Medium	express	Web server framework	20
Medium	semver	Version parsing and range handling	20
Medium	postcss	CSS transformation framework	20
Large	lodash	General-purpose utility library	20
Large	axios	HTTP client	20
Large	zod	Schema validation	20
<i>Total</i>	<i>9 dependencies</i>		<i>180</i>

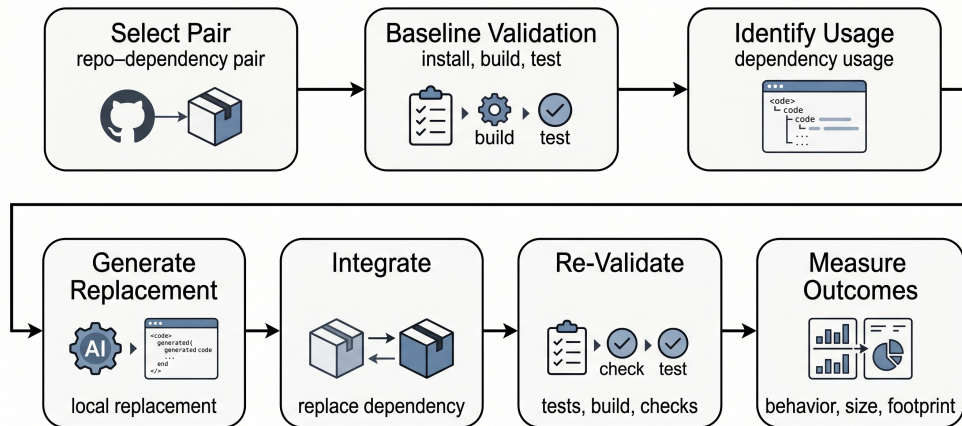
For each dependency, we select 20 evaluable client repositories (180 pairs total). We identified candidate repositories through a code search over public GitHub repositories that imported the target dependencies. This initial pool was filtered to retain only repositories suitable for evaluation: they had to build in a clean environment, expose validation artifacts (such as test suites or executable scripts), and pass their baseline validation prior to any regeneration attempt. When multiple repositories were equally suitable, repository popularity

was used as a secondary preference signal. The primary unit of analysis remains the repository–dependency pair, as regeneration feasibility depends not only on the dependency itself but also on how deeply a particular repository uses it.

4.3 Regeneration Procedure

The experimental workflow is summarized in Figure 1. For each of the 180 pairs, we execute the following phases:

1. **Establish Baseline:** We clone the repository at a fixed revision, install its declared dependencies, and run its validation artifacts (e.g., `npm test`). Only repositories with successful baseline validation are retained, providing a strict behavioral denominator for comparison.
2. **Regenerate Agentically:** First, the agent performs *usage identification* by statically extracting all first-party imports of the target dependency to scope the required functionality. Then, the agent is provided a zero-shot prompt template that is dynamically injected with the repository path and target dependency. The agent operates with internet access and can directly inspect the original upstream dependency’s source code within the local environment to guide its synthesis. It is permitted a maximum iteration budget of two attempts (one initial generation and one retry) to resolve any validation failures before terminating.
3. **Re-Validate & Measure:** This step represents the final evaluation. If the agent exhausts its iteration budget in Step 2 without satisfying the test suite, this step records a “Failed” outcome. If the final generated replacement passes all baseline checks without introducing new failures, it is recorded as a “perfect” preservation of repository-observed behavior. (Note: In Table 3, the baseline and after numbers reflect the raw count of individual assertions and tests passing). Finally, we measure the footprint differences of the successful replacements.



■ **Figure 1** Our experimental workflow. For implementation, see our artifact.

5 Results

This section reports results from our evaluation of use-case-oriented regeneration across 180 repository–dependency pairs. We interpret these results as a characterization of the feasibility and limits of repository-centered regeneration.

■ **Table 3** Summary of regeneration outcomes across 180 repository–dependency pairs. “Baseline” and “After” report the raw count of individual validation checks (e.g., assertions and tests) passing before and after regeneration. “Perfect” reports the number of repository–dependency pairs for which all baseline checks still pass after regeneration.

Dependency	Repos	Baseline	After	Failed	Perfect
<code>nanoid</code>	20	2,118	2,118	0	20/20
<code>change-case</code>	20	20,917	20,880	37	19/20
<code>chalk</code>	20	4,716	4,715	1	19/20
<code>express</code>	20	1,170	1,164	6	18/20
<code>semver</code>	20	5,111	5,109	2	18/20
<code>postcss</code>	20	2,292	2,289	3	18/20
<code>lodash</code>	20	12,804	12,700	104	16/20
<code>axios</code>	20	7,513	7,509	4	18/20
<code>zod</code>	20	7,993	7,993	0	20/20
<i>Total</i>	<i>180</i>	<i>64,634</i>	<i>64,477</i>	<i>157</i>	<i>166/180</i>

5.1 RQ1 & RQ2: Feasibility and Behavioral Preservation

As detailed in Table 3, use-case-oriented regeneration proved feasible and highly behavior-preserving across a substantial fraction of the evaluated pairs. In aggregate, generated replacements achieved a 99.8% validation pass rate, with 166 of the 180 pairs demonstrating perfect behavioral preservation. The agent identified first-party call sites, generated local replacement code, updated callers, and removed runtime reliance on the original packages.

However, outcomes varied by dependency. While highly bounded utilities like `nanoid` and `zod` saw perfect preservation across all client repositories, more complex or deeply integrated dependencies like `lodash` presented greater challenges. This variation confirms that feasibility depends heavily on both the target dependency’s complexity and how specifically the downstream repository uses it.

Crucially, this high rate of preservation must be interpreted carefully. Because preservation is measured using each repository’s existing validation artifacts, these results confirm that regeneration can seamlessly preserve *repository-observed* behavior. They do not prove full semantic equivalence across all possible inputs, as a generated replacement may pass available tests while omitting untested edge cases or undocumented compatibility behaviors. While preservation rates are high, the strength of the underlying validation artifacts varied. For instance, the average line coverage of the generated replacements executed by the repositories’ existing test suites ranged from 92.5% for utilities like `change-case` to 40.8% for complex libraries like `lodash`, underscoring that passing existing tests does not guarantee all generated logic is strictly verified.

5.2 RQ3: Size and Dependency Footprint

Successful replacements were substantially narrower than their upstream counterparts. As shown in Table 4, the generated implementations exposed dramatically smaller interfaces – averaging a 93.1% reduction in exported API surface across the benchmark.

This reduction supports the central premise of use-case-oriented regeneration: downstream repositories typically rely on only a narrow slice of a dependency’s behavior. To illustrate this footprint reduction, Figure 2 shows a generated replacement for `nanoid`.

■ **Table 4** API-surface reduction in generated replacements. The generated implementations expose substantially smaller interfaces than the original dependencies, reflecting repository-specific usage rather than full-package replacement. “Original exports” refers to the total number of publicly accessible functions, classes, and variables exposed by the upstream package’s main entry point.

Dependency	Original exports	Avg. regenerated	Reduction	Utilization
zod	236	4.1	98.3%	1.7%
lodash	306	19.4	93.7%	6.3%
chalk	64	5.0	92.2%	7.8%
postcss	25	1.0	96.0%	4.0%
change-case	14	0.4	97.1%	2.9%
axios	33	6.4	80.6%	19.4%
semver	45	9.8	78.1%	21.9%
nanoid	5	1.4	73.0%	27.0%
express	11	3.2	70.9%	29.1%
<i>Overall</i>	<i>82.1</i>	<i>5.6</i>	<i>93.1%</i>	<i>6.9%</i>

```
const urlAlphabet =
  'useandom-26T198340PX75pxJACKVERYMINDBUSHWOLF '
  + '_GQZbfgjhjklqvwyzriect';

export function nanoid(size = 21) {
  const bytes = new Uint8Array(size);
  crypto.getRandomValues(bytes);

  let id = '';
  for (let i = 0; i < size; i++) {
    id += urlAlphabet[bytes[i] & 63];
  }

  return id;
}
```

■ **Figure 2** Use-case-regenerated replacement for nanoid.

The agent successfully extracted only the core random-byte generation logic required by the client, shedding the original package’s unused CLI tools and custom alphabets. This reduces the dependent’s unused API surface and transitive exposure.

Footprint reduction alone is an imperfect proxy for quality. A smaller replacement is easier to audit, but may omit behaviors that future repository evolution might require. The primary empirical signal here is that these replacements are significantly smaller *while still preserving the current validation behavior*, suggesting regeneration is most effective for narrow, stable, and well-tested usage slices.

5.3 Failure Modes and Limitations

Analyzing the 14 failed regeneration attempts clarifies the current limits of this approach. Failures strongly correlated with specific dependency characteristics and usage breadth, typically emerging when the target dependency provided broad, subtle, or highly dynamic behaviors that were difficult for the agent to infer strictly from static call sites. Specifically, we observed three primary failure modes:

1. *Semantic and edge-case mismatches*, such as failing to reproduce **semver**’s prerelease logic or **lodash**’s deep cloning and chained APIs. These failures occurred when the repository’s usage breadth involved complex internal state that exceeded the validation suite’s explicit coverage;

2. *Class identity mismatches*, where generated objects failed strict `instanceof` checks against native or cross-package prototypes (e.g., in `zod` and `semver`); and
3. *Deep framework integrations*, where replacements could not satisfy complex internal tooling expectations, such as `express` middleware arity or `axios` mock interceptors.

In these scenarios, the agent successfully handled the common path but failed on nuanced edge cases exercised by the repository’s test suite. These failures may also reflect the limitations of our validation setup. Repository validation artifacts can be brittle, environment-sensitive, or uneven in coverage. Passing a weak test suite might hide significant semantic deviations, while failing a flaky or snapshot-sensitive test might falsely indicate a broken replacement.

Detailed descriptions of these attempts are given in our accompanying artifact [29].

6 Discussion and Research Agenda

The results indicate that use-case-oriented regeneration is a potentially viable sourcing strategy for bounded slices of functionality. However, safely realizing this vision requires shifting the research focus from package-level synthesis to automated assurance and secure agent governance.

6.1 The Security of Generated Code

The central challenge for regeneration is behavioral confidence and security. Passing existing repository tests does not guarantee safety across unexercised edge cases, nor does it prevent the introduction of AI-generated vulnerabilities or hallucinations [23]. Regeneration changes the threat model: it replaces the risk of external supply-chain vulnerabilities with the risk of local AI-generated defects. Future work must prioritize the automated security verification of the generated slice, moving beyond existing test suites to incorporate proactive AI-assisted vulnerability scanning (similar to the deep dependency threat detection pioneered by platforms like Socket [31]), as well as coverage-guided validation and differential testing against the original dependency’s behavior.

6.2 From Human Trust to Automated Assurance

Under conventional reuse, downstream projects rely on upstream maintainer reputation as a proxy for trust. Regeneration demands stronger, localized evidence. Simply expecting downstream developers to manually judge whether a dependency is narrow or stable enough for safe regeneration is insufficient and scales poorly. The research community must develop automated usage analysis and slice characterization tools to systematically determine regeneration suitability.

Furthermore, establishing trust in generative systems requires rigorous safety benchmarking for AI coding agents. The community needs standardized evaluation frameworks, such as DepDec-Bench [28], to systematically measure how safely, securely, and accurately different large language models handle dependency decomposition and synthesis before they are deployed in production environments.

For practitioners considering this approach today, we propose a preliminary heuristic checklist: regeneration should only be attempted if (1) the dependency serves a narrow, bounded role; (2) the repository possesses high-coverage validation artifacts specifically exercising that functionality; and (3) the downstream team is equipped to independently monitor the upstream package for relevant security advisories. If these conditions are not met, maintaining the standard upstream dependency remains the safer choice.

6.3 Governing AI Agents in the Software Supply Chain

Regenerated replacements capture a repository-specific slice of functionality at a single point in time, meaning they do not automatically inherit future upstream security patches or API clarifications. To manage this responsibly, we must develop secure regeneration workflows and centralized governance infrastructure.

Future software ecosystems will require a governance layer, such as AgentHub [22], that audits agent actions, manages the precise provenance of synthesized code, automatically monitors upstream changes for relevant security fixes, and enforces local security policies before any generated code is merged into the codebase. This infrastructure will shift the software supply chain from rigid upstream adoption to flexible, verifiable, and strictly governed repository-specific generation.

6.4 Implications for Engineering Processes

Use-case-oriented regeneration changes how downstream projects benefit from the “many eyes” of open-source reuse [27]. In conventional artifact-centric reuse, a widely used upstream package accumulates not only implementation fixes, but also specification knowledge: API conventions, clarified edge-case behavior, documentation, tests, examples, and compatibility expectations. These are precisely the properties that afford regeneration. A downstream project that continues to depend on the package can inherit this evolving specification through dependency updates. A regenerated replacement, by contrast, captures a repository-specific slice of that specification at generation time. It does not automatically inherit later semantic changes, security fixes, or API clarifications. Regeneration therefore entails an explicit refresh practice: monitoring upstream changes, deciding whether they affect the local API slice, and acting when they do. This situation demands regeneration provenance metadata. Without such metadata, maintainers cannot reliably determine whether a later upstream change is irrelevant, requires regeneration, or exposes an unsound local approximation.

7 Threats to Validity

As this is a vision paper with a limited measurement, several threats affect our findings:

- **Construct Validity:** Preservation is measured via existing repository tests, not formal proofs. We mitigate this by bounding claims strictly to *repository-observed* equivalence.
- **Internal Validity:** Outcomes rely on a specific agentic workflow (Claude Opus) and prompt. We isolated agent impact by requiring strict baseline pre-validation.
- **External Validity:** Evaluation is limited to JS/TS and to 9 dependencies. Feasibility may differ in ecosystems with different testing cultures or agent training data.

8 Conclusion

Modern software systems rely on complex dependency graphs with substantial ongoing maintenance costs. We argue that recent advances in generative AI warrant reexamining the traditional build-*vs.*-reuse tradeoff. We envision *use-case-oriented regeneration* as a strategy in which generative systems synthesize only the repository-specific slice of an upstream dependency that is actually required. Our evaluation of 180 repository-dependency pairs suggests that this approach is often technically feasible and can substantially reduce dependency footprint while preserving repository-observed behavior.

Whether regeneration ultimately proves preferable to conventional reuse in a given context will depend on more than implementation cost. Regenerated code forfeits some of the advantages of shared upstream artifacts, including continual optimization, community review, evolving specifications, and Raymond’s “many eyes” effect that has historically improved mature open-source software. Conversely, it offers opportunities for repository-specific optimization, reduced attack surface, simplified auditing, and greater local control. Determining when these competing considerations favor regeneration over continued upstream dependence remains an important open research question. We therefore view use-case-oriented regeneration not as a replacement for software package ecosystems, but as a complementary sourcing strategy. Future work should develop principled methods for deciding when regeneration is appropriate, how regenerated implementations should be validated and refreshed, and how engineers can balance the long-term benefits of shared ecosystems against the flexibility of locally synthesized software. Ultimately, use-case-oriented regeneration has the potential to shift the software supply chain from rigid upstream adoption to more flexible, verifiable, and repository-specific generation, granting downstream projects greater control over the code they trust and maintain.

References

- 1 Rabe Abdalkareem, Vinicius Oda, Suhaib Mujahid, and Emad Shihab. On the impact of using trivial packages: an empirical case study on npm and pypi. *Empirical Software Engineering*, 25:1168–1204, 2020. doi:10.1007/S10664-019-09792-9.
- 2 Paschal C. Amusuo, Kyle A. Robinson, Santiago Torres-Arias, Laurent Simon, and James C. Davis. Ztd_java: Mitigating software supply chain vulnerabilities via zero-trust dependencies. In *International Conference on Software Engineering (ICSE)*, 2025.
- 3 Barry W. Boehm. Software engineering economics. *IEEE transactions on Software Engineering*, 10(1):4–21, 1984. doi:10.1109/TSE.1984.5010193.
- 4 James C. Davis. Model-based agentic software engineering. <https://davisjam.github.io/model-based-agentic-software-engineering/book/mage-book.pdf>, 2026. Online book. Accessed: 2026-08-05.
- 5 James C Davis, Paschal C Amusuo, Tanmay Singla, Berk Çakar, and Kirsten A Davis. Cheap code, costly judgment: A case study on governable agentic software engineering. [*arXiv’26/arXiv preprint*, 2026. arXiv:26].
- 6 Sivana Hamer, Nasif Imtiaz, Mahzabin Tamanna, Preya Shabrina, and Laurie Williams. Trusting code in the wild: Exploring contributor reputation measures to review dependencies in the rust ecosystem. *IEEE Transactions on Software Engineering*, 2025.
- 7 Ruiyin He, Tomi Saarinen, Nyyti Saarinen, Arie van Deursen, and Mika Mäntylä. Automating dependency updates in practice: An exploratory study on security vulnerabilities and other dependency update issues. *IEEE Transactions on Software Engineering*, 2023.
- 8 Fang Hou and Slinger Jansen. A systematic literature review on trust in the software ecosystem. *Empirical Software Engineering*, 28(1), 2023. doi:10.1007/S10664-022-10238-Y.
- 9 Abbas Javan Jafari, Diego Elias Costa, Emad Shihab, and Rabe Abdalkareem. Dependency update strategies and package characteristics, 2023. doi:10.48550/arXiv.2305.15675.
- 10 Dhanushka Jayasuriya, Samuel Ou, Saakshi Hegde, Valerio Terragni, Jens Dietrich, and Kelly Blincoe. An extended study of syntactic breaking changes in the wild. *Empirical Software Engineering*, 30(2), 2024. doi:10.1007/S10664-024-10563-4.
- 11 Carlos E. Jimenez, John Yang, Alexander Wettig, Shunyu Yao, Kexin Pei, Ofir Press, and Karthik Narasimhan. Swe-bench: Can language models resolve real-world github issues? *arXiv preprint*, 2023. doi:10.48550/arXiv.2310.06770.

- 12 Kelechi G. Kalu, Sofia Okorafor, Betül Durak, Kim Laine, Radames C. Moreno, Santiago Torres-Arias, and James C. Davis. Arms: A vision for actor reputation metric systems in the open-source software supply chain. In *Journal Ahead Workshop (JAWs)*, 2026.
- 13 Kelechi G. Kalu, Tanmay Singla, Chinenye Okafor, Santiago Torres-Arias, and James C. Davis. An industry interview study of software signing for supply chain security. In *USENIX Security Symposium*, pages 81–100, 2025. URL: <https://www.usenix.org/conference/usenixsecurity25/presentation/kalu>.
- 14 Evangelos Lamprou, Julian Dai, Grigoris Ntousakis, Martin C. Rinard, and Nikos Vasilakis. Lexo: Eliminating stealthy supply-chain attacks via llm-assisted program regeneration, 2025. doi:10.48550/arXiv.2510.14522.
- 15 Enrique Larios Vargas, Maurício Aniche, Christoph Treude, Magiel Bruntink, and Georgios Gousios. Selecting third-party libraries: the practitioners’ perspective. In *Foundations of Software Engineering (ESEC/FSE)*, 2020.
- 16 Linux Foundation. A summary of census ii: Open source software application libraries the world depends on, 2022. URL: <https://www.linuxfoundation.org/blog/blog/a-summary-of-census-ii-open-source-software-application-libraries-the-world-depends-on>.
- 17 Yuxin Liu, Deepika Tiwari, Cristian Bogdan, and Benoit Baudry. Detecting and removing bloated dependencies in commonjs packages. *Journal of Systems and Software*, 230:112509, 2025. doi:10.1016/J.JSS.2025.112509.
- 18 Ivano Malavolta, Kishan Nirghin, Gian Luca Scoccia, Simone Romano, Salvatore Lombardi, Giuseppe Scanniello, and Patricia Lago. Javascript dead code identification, elimination, and empirical assessment. *IEEE Transactions on Software Engineering*, 49(7):3692–3714, 2023. doi:10.1109/TSE.2023.3267848.
- 19 Mark Marron. Toward an agentic infused software ecosystem, 2026. doi:10.48550/arXiv.2602.20979.
- 20 Marc Ohm, Henrik Plate, Arnold Sykosch, and Michael Meier. Backstabber’s knife collection: A review of open source software supply chain attacks. In *Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*, pages 23–43, 2020. doi:10.1007/978-3-030-52683-2_2.
- 21 Chinenye Okafor, Taylor R. Schorlemmer, Santiago Torres-Arias, and James C. Davis. Sok: Analysis of software supply chain security by establishing secure design properties. In *Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (SCORED)*, pages 15–24, 2022. doi:10.1145/3560835.3564556.
- 22 Erik Pautsch, Tanmay Singla, Parv Kumar, Wenxin Jiang, Huiyun Peng, Behnaz Hassanshahi, Konstantin Läufer, George K. Thiruvathukal, and James C. Davis. AgentHub: A registry for discoverable, verifiable, and reproducible ai agents. In *The Journal Ahead Workshop (JAWs 2026)*, co-located with the 48th International Conference on Software Engineering (ICSE 2026), April 2026.
- 23 Hammond Pearce, Baleegh Ahmad, Benjamin Tan, Brendan Dolan-Gavitt, and Ramesh Karri. Asleep at the keyboard? assessing the security of github copilot’s code contributions, 2021. arXiv:2108.09293.
- 24 pip developers. Vendoring policy. URL: <https://pip.pypa.io/en/latest/development/vendoring-policy/>.
- 25 Serena Elisa Ponta, Wolfram Fischer, Henrik Plate, and Antonino Sabetta. The used, the bloated, and the vulnerable: Reducing the attack surface of an industrial application. In *International Conference on Software Maintenance and Evolution (ICSME)*, pages 555–558, 2021. doi:10.1109/ICSME52107.2021.00056.
- 26 Anh Quach, Hoa Nguyen, Derek Bruening, Jason Hiser, Cristiano Giuffrida, Vasileios P. Kemerlis, Ramesh Peri, and Michael Payer. Debloating software through piece-wise compilation and loading. In *USENIX Security Symposium*, pages 341–356, 2018.
- 27 Eric Raymond. The cathedral and the bazaar. *Knowledge, Technology & Policy*, 12(3):23–49, 1999.

- 28 Tanmay Singla, Berk Çakar, Paschal C. Amusuo, and James C. Davis. Towards a benchmark for dependency decision-making, 2026. doi:10.48550/arXiv.2601.00205.
- 29 Tanmay Singla and James C. Davis. Software supply chains are dead: Use-case-oriented regeneration, May 2026. doi:10.5281/zenodo.20370439.
- 30 SLSA Contributors. Supply-chain levels for software artifacts (slsa) provenance specification, 2023. URL: <https://slsa.dev/spec/v1.0/provenance>.
- 31 Socket. Socket security: Proactive open source supply chain protection. <https://socket.dev/>, 2026.
- 32 Sonatype. 2026 state of the software supply chain report, 2026. URL: <https://www.sonatype.com/state-of-the-software-supply-chain>.
- 33 Synopsys Cybersecurity Research Center. Open source security and risk analysis (OSSRA) report. Technical report, Synopsys, Inc., 2024.
- 34 Tamanna, Greg Tystahl, Nusrat Zahan, Patrick Morrison, Yasemin Acar, Michel Cukier, Christian Kästner, Alexandros Kapravelos, Dominik Wermke, and William Enck. Research directions in software supply chain security. *ACM Transactions on Software Engineering and Methodology*, 34(5), 2025. doi:10.1145/3714464.
- 35 Nikos Vasilakis, Achilles Benetopoulos, Shivam Handa, Alizee Schoen, Jiasi Shen, and Martin C. Rinard. Supply-chain vulnerability elimination via active learning and regeneration. In *Conference on Computer and Communications Security (CCS)*, pages 1755–1770, 2021. doi:10.1145/3460120.3484736.
- 36 Qian Wang et al. OpenHands: An open platform for ai software developers as generalist agents, 2024. URL: <https://github.com/OpenHands/OpenHands>.
- 37 webpack. Tree shaking, 2025. URL: <https://webpack.js.org/guides/tree-shaking/>.
- 38 N. R. Weeraddana, Mahmoud Alfadel, and Shane McIntosh. Dependency-induced waste in continuous integration: An empirical study of unused dependencies in the npm ecosystem. *ACM Transactions on Software Engineering and Methodology*, 2024.
- 39 Erik Wittern, Philippe Suter, and Shriram Rajagopalan. A look at the dynamics of the javascript package ecosystem. In *Mining Software Repositories (MSR)*, pages 351–361, 2016. doi:10.1145/2901739.2901743.
- 40 John Yang, Yu Zhang, Zihang Chen, et al. Swe-agent: Agent-computer interfaces enable automated software engineering. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- 41 Markus Zimmermann, Cristian-Alexandru Staicu, Cam Tenny, and Michael Pradel. Small world with high risks: A study of security threats in the npm ecosystem. In *USENIX Security Symposium*, pages 995–1010, 2019. URL: <https://www.usenix.org/conference/usenixsecurity19/presentation/zimmerman>.