

SELF-AUDIT RESULTS

Case Study: Autonomous Codebase Audit

25-file Python codebase — 693 tests, zero regressions

THE CHALLENGE

Smelt was pointed at its own source code as a stress test: 25 Python files comprising the analysis engine, patch system, MCP server, CLI, and tests. The audit needed to find real issues that manual review had missed — then fix them without breaking any of the 693 existing tests.

WHAT WAS FOUND

The audit identified **16 critical issues** and **2 logic bugs**, organized into 6 prioritized batches by severity class:

BATCH	CATEGORY	FOUND	STATUS
1 Logic bugs	Off-by-one, wrong return	2	Already fixed
2 Crash paths	Unguarded JSON parsing	9	8 done, 1 new fix
3 Silent swallowing	Bare except, no logging	4	3 done, 1 improved
4 Type safety	SgNode missing annotations	6	3 done, 3 annotated
5 Bare dict types	Untyped dict parameters	6	4 done, 2 fixed
6 Standards	Long functions, deep nesting	12	10 done, 2 extracted

HOW IT WORKED

Each batch was applied sequentially with full test-suite validation between rounds. Fixes ranged from type annotations and log improvements to function extraction and JSON decode guards. The system identified that many issues were already addressed — avoiding redundant changes.

KEY INSIGHT

73% of flagged issues were already resolved, demonstrating the cumulative effect of iterative quality passes. The remaining 6 fixes each targeted a specific defect class.

16

Critical issues identified across 6 defect categories

6

New fixes applied — targeted, surgical corrections

0

Regressions introduced 693 of 693 tests passing

25

Python source files audited end-to-end

73%

Of issues already resolved from prior quality passes

6

Prioritized batches — logic → crashes → types → style

BEFORE AUDIT

16 issues

2 logic bugs, 9 crash paths, 4 silent exceptions, systematic type & standards violations

AFTER AUDIT

0 remaining

6 new fixes applied, 693/693 tests green, zero regressions across all batches

"The tool audited its own codebase, found real bugs, fixed them in priority order, and never broke a single test. That's the bar."

— Smelt self-audit, March 2026

