The State of AI Tool and Agent Sprawl, 2026

An 890-Target Publication Subset on Public Agent Declarations, Approval Visibility, and Governance Evidence

- Report ID: `ai-tool-sprawl-v2-2026`
- Version: `v2-subset-890`
- Run ID: `sprawl-v2-full-20260312b`
- Subset basis: completed targets from the frozen 1,000-target publication cohort (`890/1000`)
- Campaign design: `deterministic baseline`, public-repository cohort, one repo per owner, clone-sourced scan artifacts
- Wrkr commit pin: `fb819fb9ff235f1160c21103779904f871918cf3` (`v1.0.8`)
- Prepared by Centre for AI Security and Integrity (CAISI). Contact: research@caisi.dev. Full artifacts: github.com/Clyra-AI/safety.

Authorship and Affiliation

David Ahmann - Head of Cloud, Data and AI Platforms at CDW Canada (LinkedIn) Talgat Ryshmanov - Principal DevSecOps Consultant at Adaptavist (LinkedIn)

About CAISI

The Centre for AI Security and Integrity (CAISI) publishes independent, reproducible research on AI governance. Every headline number in this manuscript maps to machine-generated artifacts and deterministic queries in the repository.

This report is narrower than a market census. It measures what a frozen public GitHub cohort exposes about AI tools, agent declarations, approval posture, evidence readiness, and control-aligned artifacts. It does not claim visibility into private repositories, production credentials, or internal runtime paths.

Executive Summary

This manuscript is tied to the **890 completed targets** from a frozen **1,000-target** publication cohort. The original run stopped before all targets finished after repeated disk-budget exhaustion. Rather than infer the missing `110`, this report uses only the completed subset and labels it explicitly as such.

The strongest signal in the subset is not privileged agent execution. It is governance opacity around AI adoption in software-delivery environments. In the measured subset, **874 of 890 targets (98.2%)** declared at least one agent, but only **36 of 890 (4.04%)** showed any deployable-agent signal, and **100% of detected agents** were missing at least one declared binding. At the same time, **419 of 890 targets (47.08%)** lacked verifiable governance evidence in the repository, and non-source AI tools outside the baseline-approved set outnumbered baseline-approved tools **11:1**.

The public compliance signal is similarly shallow. Across `889` rolled-up targets, EU AI Act proxy coverage averaged **33.33%**, effectively one of three current proxy controls. SOC 2 and PCI DSS proxy coverage were **0%** in the public artifact set. These are visibility findings, not legal conclusions.

One caution matters for interpretation. The unfinished `110` targets are all from the AI-native stratum. Relative to the planned `50/30/20` publication mix, the completed subset is slightly underweighted on AI-native projects and therefore likely understates AI-native-heavy signals such as transparency gaps, approval opacity, and weak evidence posture. It may also slightly overstate deployable-agent prevalence because the completed dev-platform share is comparatively larger.

Key Findings (At a Glance)

- Targets with declared agents: `98.2%` (`874/890`)

- Average declared agents per target: 8.97
- Targets with deployed-agent signal: 4.04% (36/890)
- Declared agents with incomplete bindings: 100% (7984/7984)
- Targets without verifiable governance evidence: 47.08% (419/890)
- Not-baseline-approved to baseline-approved tool ratio: 11:1 (715/65)
- Targets with Article 50 transparency-proxy gap: 17.19% (153/890)
- Targets with write-capable agents detected in public artifacts: 0%

Headline Integrity Block

All headline claims in this manuscript map to one immutable partial-subset artifact package.

- Run ID: sprawl-v2-full-20260312b
- Subset target file:

runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/completed-targets-890.md

- Claim artifact:

runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/claims-finalized-v2.json

- Aggregate artifact:

runs/tool-sprawl/sprawl-v2-full-20260312b/agg/campaign-summary-v2.json

Key	Headline number	Denominator
H1	11	baseline-approved non-source tools
H2	98.2	targets scanned
H3	8.97	targets scanned
H4	4.04	targets scanned
H5	100	declared agents
H6	47.08	targets scanned
H7	17.19	targets scanned
H8	890	completed subset

Claim-key map:

- H1 = sprawl_v2_not_baseline_approved_to_approved_ratio
- H2 = sprawl_v2_orgs_with_agents_pct
- H3 = sprawl_v2_avg_agents_per_org
- H4 = sprawl_v2_orgs_with_deployed_agents_pct
- H5 = sprawl_v2_agents_missing_bindings_pct
- H6 = sprawl_v2_orgs_without_verifiable_evidence_pct
- H7 = sprawl_v2_article50_gap_prevalence_pct
- H8 = sprawl_v2_orgs_scanned

Deterministic query map:

```
H1  jq '.campaign.metrics.not_baseline_approved_to_approved_ratio'
H2  jq '.campaign.metrics.orgs_with_agents_pct'
H3  jq '.campaign.metrics.avg_agents_per_org'
H4  jq '.campaign.metrics.orgs_with_deployed_agents_pct'
H5  jq '.campaign.metrics.agents_missing_bindings_pct'
H6  jq '.campaign.metrics.orgs_without_verifiable_evidence_pct'
H7  jq '.campaign.metrics.article50_gap_prevalence_pct'
H8  jq '.campaign.metrics.orgs_scanned'
```

Additional supporting claims, including write-capable, exec-capable, and attack-path prevalence, are mapped in runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/claims-finalized-v2.json.

1) What We Scanned

The frozen publication cohort for this report originally contained 1,000 public owner/repo targets selected under the v2 publication profile with one repository per owner. The collection run completed 890 targets before stopping. This manuscript is tied only to those completed targets and does not extrapolate the unfinished 110.

Completed cohort composition:

Cohort	Targets completed
AI-native	390
Dev platform	300
Security platform	200
Total	890

The finished subset preserves the full dev-platform and security-platform strata. The missing 110 targets are all AI-native. That skews the final subset from the planned 50/30/20 mix toward 43.8/33.7/22.5.

Headline tool counts exclude tool_type == "source_repo". That matters because raw tool detections were dominated by repository-native rows.

Measure	Value
Targets scanned	890
Raw detected tools	7927
source_repo rows excluded from headline scope	7147
Detected non-source tools	780
Targets with at least one detected non-source tool	166
Targets with at least one declared agent	874
Declared agents	7984
Targets with at least one deployed agent	36

This report therefore measures a very broad declaration surface, but a much narrower verifiable deployment surface.

2) Why This Matters for Software Delivery and AppSec

Public AI adoption signals matter most when they intersect software-delivery paths. A repository that declares agents, references model-facing tooling, or encodes orchestration patterns is not automatically risky. The AppSec question is narrower: can an organization show which AI paths can influence code, CI, workflow, or operational delivery, and can it prove the approval and evidence chain around those paths?

In this subset, the measurable public signal is that AI and agent declarations are common, but binding completeness, approval visibility, and control-aligned evidence are weak. That gap is operationally important even when privileged runtime paths are not visible. When security or engineering leadership asks whether AI-assisted delivery is governed, the burden is not only to answer ”yes” in principle. It is to produce machine-verifiable evidence for the specific tools and agent surfaces in scope.

This is why the main story in this dataset is evidence posture, not exploitability. Public repositories rarely expose production credentials, runtime identities, or live approval workflows. They do, however, expose whether organizations appear ready to explain and evidence AI use in delivery-adjacent codebases. On that question, the subset shows persistent gaps.

3) Tool and Agent Adoption Signal

The completed subset shows near-universal agent declaration. 874/890 targets (98.2%) exposed at least one agent row, and the average target exposed 8.97 declared agents. That is a large declared surface relative to the headline non-source tool inventory, which totaled 780 non-source tools across only 166 targets.

The tool approval picture is weaker than the raw adoption signal. In headline scope:

Tool approval posture	Count
Baseline-approved tools	65
Explicit-unapproved tools	0
Approval-unknown tools	715
Not-baseline-approved tools	715
Not-baseline-approved to baseline-approved ratio	11:1

At the target level:

- 166/890 targets (18.65%) had at least one detected non-source tool.
- 53/890 (5.96%) had at least one baseline-approved non-source tool.
- 135/890 (15.17%) had at least one not-baseline-approved non-source tool.

The adoption mix varies by cohort. AI-native targets were the most tool-heavy in headline scope: 30.0% of completed AI-native targets had at least one non-source tool, versus 11.67% in dev-platform and 7.0% in security-platform. Dev-platform targets showed the highest deployable-agent share at 6.67%, while security-platform targets showed the lowest agent density overall.

The raw inventory therefore supports two conclusions at once. First, public AI and agent declarations are common across the cohort. Second, most of the concrete, non-source tool signal is not paired with deterministic baseline approval evidence.

4) Delivery Surface Exposure

The delivery-surface signal in this subset is wide in declaration terms and thin in operational completeness terms.

Exposure qualifier	Value
Targets with declared agents	874
Targets with deployed-agent signal	36
Deployed agents detected	65
Binding-complete agents	0
Binding-incomplete agents	7984
Targets with write-capable agents	0
Targets with exec-capable agents	0
Targets with agent-linked attack paths	0

The most important line in that table is the binding result. Every detected agent was missing at least one declared binding, and 98.2% of targets had at least one binding-incomplete agent. In plain terms, public repositories often show that an agent concept exists, but usually do not show enough about tool bindings, data bindings, or auth bindings to treat the agent as a fully evidenced delivery actor.

That is also why the zero values for write-capable agents, exec-capable agents, and attack-path prevalence should be read carefully. The public subset does not show those capabilities. It does not prove they are absent in private runtime environments. This dataset is strongest at showing where declaration outruns evidence, not at proving internal privilege is low.

5) Governance and Evidence Gaps

The strongest governance finding in the subset is the evidence gap.

Evidence posture	Count
`verifiable` evidence tier	471
`basic` evidence tier	419
Targets without verifiable evidence	419

That yields `47.08%` of the completed subset without verifiable evidence. Put differently, nearly half of the scanned targets exposed enough AI or agent signal to be in scope but not enough evidence to clear the deterministic `verifiable` threshold.

Approval visibility is also weak. There were no explicit-unapproval markers in headline scope, but the absence of negative markers is not a positive governance result. The measurable issue is that the system could not deterministically establish approval for most detected tools. The not-baseline-approved to baseline-approved ratio of `11:1` is therefore better understood as an approval-proof gap than as a danger score.

The cohort split reinforces that interpretation. Targets without verifiable evidence were:

- `53.08%` in AI-native
- `49.33%` in dev-platform
- `32.0%` in security-platform

Security-platform repositories were measurably better on public evidence posture, but even there roughly one third failed to expose verifiable evidence under the current baseline. AI-native and dev-platform targets were worse, and the missing `110` AI-native targets likely make the current `47.08%` a conservative subset estimate.

6) Regulatory Readiness

Regulatory outputs in this report are deterministic evidence-of-control proxies. They are not legal or audit conclusions.

The subset showed `153/890` targets (`17.19%`) with an Article 50 transparency-proxy gap. Cohort differences were meaningful:

- AI-native: `28.21%`
- Dev-platform: `10.0%`
- Security-platform: `6.5%`

Across `889` rolled-up targets with framework output:

Framework family	Average proxy coverage	Readout
EU AI Act	33.33%	effectively `1 of 3` proxy controls visible
SOC 2	0%	no visible proxy coverage in public artifacts
PCI DSS	0%	no visible proxy coverage in public artifacts

The EU AI Act result is not broad compliance readiness. It means the public artifact set consistently exposed a minimal governance floor while leaving two of the three current proxy controls unproven. The SOC 2 and PCI DSS results are even narrower: the deterministic public artifact set did not provide machine-verifiable proxy coverage for those control families in the completed subset.

For a software-delivery and AppSec audience, that is the meaningful conclusion. Public AI and agent signals were common, but the evidence needed to support mature control assertions was sparse.

7) Case Studies

This section uses anonymized examples drawn from completed appendix rows. They are illustrative patterns, not headline cases.

- Example A, AI-native framework repository: 128 declared agents, 124 not-baseline-approved non-source tools, 0 approved tools, article50_gap=true, evidence_tier=basic. This is the clearest expression of declaration-heavy adoption with weak public approval proof.
- Example B, AI-native runtime repository: 40 declared agents, 21 deployed-agent signals, 27 non-source tools, 25 not-baseline-approved tools, 2 approved tools, article50_gap=true, evidence_tier=basic. This is the strongest operational-looking example in the subset, but it still does not clear the evidence threshold.
- Example C, platform repository with stronger evidence posture: 138 declared agents, 2 deployed-agent signals, 0 non-source tools, article50_gap=false, evidence_tier=verifiable. This pattern shows that a repository can expose significant agent declaration volume without also exposing external-tool approval ambiguity.

These examples support the broader point that the public AI surface is heterogeneous. The common weakness is not one framework family or one codebase style. It is the mismatch between declaration visibility and governance evidence.

8) Methodology

The subset uses the v2 deterministic-baseline pipeline and the frozen publication cohort generated under the v2 selection profile. Source material came from out-of-box `wrkr scan --json` runs against local clones of public repositories, one repository per owner.

Key methodological facts:

- Frozen publication cohort file: internal/repos-v2-publication-1000.md
- Completed subset file used for this manuscript: runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/completed-targets-890.md
- Wrkr pin: fb819fb9ff235f1160c21103779904f871918cf3 (v1.0.8)
- Headline tool scope excludes tool_type == "source_repo"
- Framework families treated as headline-eligible in v2: EU AI Act, SOC 2, PCI DSS
- Rebuilt aggregate artifact: runs/tool-sprawl/sprawl-v2-full-20260312b/agg/campaign-summary-v2.json
- Rebuilt appendix artifact: runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/combined-appendix-v2.json

Validation status for this subset:

- relaxed validation passed with failures=0
- required threshold checks passed 6/6
- recommended threshold checks passed 5/11

The subset is internally coherent, but it is not a strict full-lane publication package. Gold-label calibration review was not completed for this partial run, and the original run did not reach its intended 1,000 completed targets.

9) Recommendations

1. Treat agent declaration inventory and deployable-agent evidence as separate metrics. A repository that declares agents is not automatically exposing a deployed agent surface.
2. Make approval evidence machine-readable. Most of the tool posture problem in this subset is unresolved approval status, not explicit disapproval.
3. Treat missing bindings as a governance finding, not a minor metadata issue. If tool, data, or auth bindings are missing, the organization cannot cleanly evidence operational boundaries for the agent.

4. Do not interpret public 0% write-capable or exec-capable agents as proof of low internal runtime risk. Public repositories systematically underexpose those paths.
5. Use proof-backed control artifacts early. EU AI Act, SOC 2, and PCI DSS proxy readiness all look weak when repositories do not expose durable evidence records.

10) Appendix

Primary appendix artifacts for this manuscript:

- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/combined-appendix-v2.json
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/tool-inventory.csv
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/agent-inventory.csv
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/agent-privilege-map.csv
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/framework-rollups.csv
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/regulatory-gap-matrix-v2.csv
- runs/tool-sprawl/sprawl-v2-full-20260312b/appendix/org-summary-v2.csv

The run-local finalized claims and threshold outputs for this manuscript are:

- runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/claims-finalized-v2.json
- runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/threshold-evaluation-v2.json

Limitations

- This manuscript is based on a completed subset (890/1000) rather than the full frozen publication cohort.
- The missing 110 targets are all AI-native, so the subset likely understates AI-native-heavy signals such as Article 50 gaps, approval opacity, and weak evidence posture.
- Public repositories understate internal runtime privilege, deployment state, credential exposure, and approval workflows.
- Headline tool metrics exclude source_repo rows by design.
- Regulatory outputs are deterministic readiness proxies, not legal determinations.

Threats to Validity

- Sample-completion bias: because the unfinished targets are all AI-native, the subset is not the exact intended publication mix.
- Visibility bias: public repository artifacts reveal less than internal systems about runtime identities and production-connected agent behavior.
- Detector-coverage bias: some frameworks and config conventions will be easier to detect than others.
- Mapping bias: control proxies depend on the currently locked rule and framework mappings.

Residual Risk

- Internal and private environments may have materially higher privilege and deployment risk than this public subset exposed.
- Targets classified as basic evidence posture may still operate significant AI or agent workflows internally without public proof artifacts.
- A public zero on write-capable, exec-capable, or attack-path metrics should not be read as a runtime safety claim.

Reproducibility Notes

- Rebuild command used for this manuscript:

```
pipelines/sprawl/rebuild_from_scans_v2.sh \
  --run-id sprawl-v2-full-20260312b \
  --targets-file runs/tool-sprawl/sprawl-v2-full-20260312b/artifacts/completed-targets-890.md \
  --mode baseline-only
```

- Claim finalization command used for this manuscript:

```
pipelines/sprawl/finalize_claims_v2.sh \
  --run-id sprawl-v2-full-20260312b \
  --lane test \
  --validate
```

- Validation command used for this manuscript:

```
pipelines/sprawl/validate_v2.sh \
  --run-id sprawl-v2-full-20260312b \
  --lane test
```

- PDF build command:

```
pipelines/common/build_report_pdf.sh --report-dir reports/ai-tool-sprawl-v2-2026
```

This subset report is strongest as a public evidence and governance-readiness measurement. It is weaker as a direct runtime privilege study. That distinction is not a weakness in presentation. It is the central empirical result.