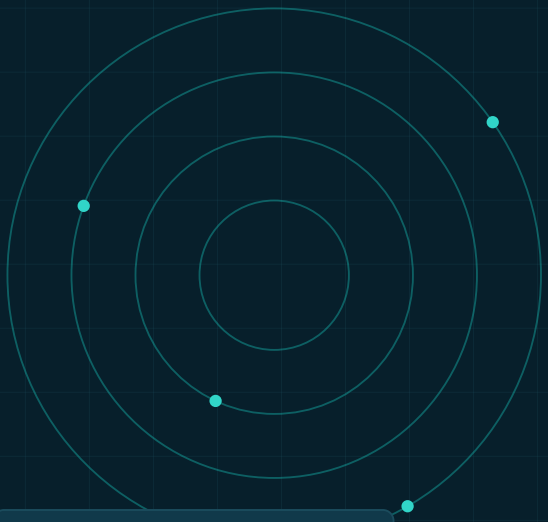
The MCP & AI Agent Trust Preflight



20 checks before you let an MCP connect or an agent act.

20 CHECKS | 4 PARTS | MCP + AGENT | LOCAL-FIRST

10

MCP CHECKS

Surface, schema, secrets, drift

6

AGENT CHECKS

Reach, chains, blast radius

4

PROOF CHECKS

Freshness, evidence, mediation

THE OPERATING PRINCIPLE

Inspect. Baseline. Re-check.

Every check can be run by hand. ECZ-ID automates the repeatable work without turning configuration evidence into a false safety verdict.

A manual preflight you can actually defend.

You would not deploy a service you had never inspected. Yet teams routinely wire MCP servers into editors or let agents act having reviewed little more than a name.

HONEST SCOPE MATTERS

These checks describe configuration, declared capability, drift, authority and evidence. They do not turn an MCP server or agent into a guaranteed-safe system.

GOOD ANSWER Specific enough that another engineer can reproduce the check and reach the same conclusion.

BAD ANSWER A reassuring label, screenshot or assumption with no baseline, no scope and no evidence trail.

The MCP surface

Before trust, inventory. You cannot reason about what you have not enumerated.

1 Enumerate every MCP server

Find workspace and user MCP configuration, legacy client files and server declarations. Record source and location.

2 Identify transport and executable

Distinguish STDIO from HTTP. Record the command basename or remote origin without leaking credentials.

3 Read tool descriptions like code

Tool descriptions influence model behaviour. Review names, descriptions and declared capabilities rather than trusting labels.

4 Inspect schemas, not just names

Look for powerful inputs, overly broad shapes and parameter changes that expand what a tool can do.

5 Name dangerous capabilities

Call out shell, filesystem, network, package, deployment, database or destructive capabilities explicitly.

6 Audit credential-shaped env names

Inventory environment variable names without exposing values. Unexpected token/key names are a review signal.

Change and drift

The server you reviewed in March is not necessarily the server you are running today.

7 Record a baseline you can compare

Capture privacy-filtered structure: transport, command shape, environment key names, tool/schema digests and origins.

8 Use fingerprints that survive secret rotation

A useful fingerprint changes when posture changes, not when an API key rotates.

9 Treat tool redefinition as high-signal

A description, schema or capability change can materially alter behaviour without changing the server name.

10 Re-check before reliance

Freshness matters. Re-run after config changes, package updates, new tools, new origins or before high-reliance use.

Agent authority

An agent risk is not what it is called. It is everything it can reach.

11 Inventory agent surfaces by content

Find manifests, tool specs, framework configs, webhook definitions and MCP relationships - not filenames alone.

12 Map real reachable capability

Follow declared tools, MCP targets, API origins and permissions to understand the actual authority graph.

13 Separate read, write and destructive power

Do not collapse file-read, external send, package mutation and deployment into one vague permission bucket.

14 Look for dangerous action chains

Credential access plus external send; filesystem plus network; package mutation plus deployment - chains matter.

15 Find credential exposure paths

Track credential-shaped key names to the capabilities that can read, transform or transmit data.

16 Make Agent-to-MCP dependence explicit

An agent can inherit the risk and authority of every MCP server it can invoke. Review both sides together.

Proof and mediation

The difference between believing and knowing is whether the claim can be re-checked.

17 Separate observed from enforced

Config inspection is OBSERVED. Runtime ENFORCED should only appear while a live mediation path actually exists.

18 Minimise secrets before runtime

When mediation is used, launch servers with only the environment keys they actually need. Do not display secret values.

19 Use deterministic local policy

For high-risk tools, explicit ALLOW / WARN / REVIEW / BLOCK / QUARANTINE rules are easier to audit than hidden judgement.

20 Demand evidence and re-check

Keep evidence you can hand to another reviewer. Re-check freshness, authority and current state before reliance.

Run the repeatable work automatically.

Twenty checks are useful as a manual discipline. ECZ-ID MCP Trust and Agent Trust turn the repeatable parts into local, reviewable workflow.

ECZ-ID MCP Trust

CONFIG, CHANGE, RUNTIME

- MCP X-Ray inventory
- Change alerts and baselines
- Credential-shaped-name warnings
- Pro: Local Trust Gate
- Pro: Secret Shield + epochs

COMMUNITY: FREE - NO ACCOUNT REQUIRED

ECZ-ID Agent Trust

AUTHORITY, REACH, CHAINS

- Agent X-Ray inventory
- Authority Graph + reachability
- Dangerous action-chain indicators
- Authority Epochs
- Reciprocal MCP view

COMMUNITY: FREE - NO ACCOUNT REQUIRED

Community gives useful evidence. Pro adds depth and control.

CAPABILITY	COMMUNITY	PRO
Inventory + declared capability	Yes	Yes
Single baseline + change detection	Yes	Yes
Full epoch history + exact diffs	-	Yes
Authority Graph + action chains	-	Yes
Secret Shield preview	-	Yes
Optional Local Trust Gate + MediationProof	-	Yes
Exportable evidence / deeper remediation	-	Yes

BOUNDARY

Community does not mediate runtime execution. Pro may optionally run a Local Trust Gate; ENFORCED appears only while the mediation evidence exists.

Start free. Add depth only where you need it.

The fastest way to improve an MCP or agent workspace is to make the invisible surface visible first - then baseline it, diff it and mediate only where the risk justifies it.

MCP TRUST PRO

£12.99

per month

£119 / year

AGENT TRUST PRO

£12.99

per month

£119 / year

DEVELOPER TRUST

£19.99

per month

£199 / year

1 INSTALL FREE — 2 SCAN / REVIEW — 3 BASELINE — 4 GET PRO — 5 ACTIVATE — 6 RE-CHECK

GET STARTED

MCP Trust	https://developers.ecocitizenz.com/mcp-trust
Agent Trust	https://developers.ecocitizenz.com/agent-trust
Developer Trust bundle	https://developers.ecocitizenz.com/developer-trust
Already purchased? Activate	https://trustops.ecocitizenz.com/developer-trust/activate

Early launch note: Pro access is designed to be available after payment confirmation. If an automated step is delayed, support can securely complete provisioning.

OPEN DEVELOPER TRUST

The 10-minute team handoff.

Use this page before a new MCP server, agent integration or material configuration change. Save the evidence with the change ticket or review record.

- **Inventory complete**
  Every server, agent surface and relevant config is named.
- **Execution path understood**
  Transport, command or remote origin is known.
- **Dangerous capabilities explicit**
  Shell, filesystem, network, write/destructive and deployment rights are visible.
- **Credential names reviewed**
  Credential-shaped environment key names are visible without exposing values.
- **Baseline captured**
  Current structure is fingerprinted and ready for comparison.
- **Drift reviewed**
  Tool, schema, origin and authority changes are understood.
- **Authority graph reviewed**
  Agent-to-tool-to-MCP reach is understood.
- **High-risk chains reviewed**
  Credential + external action, filesystem + network, package + deployment chains are visible.
- **Evidence state labelled honestly**
  Observed is not called enforced. Enforced needs current mediation proof.
- **Re-check date/trigger set**
  Know when this evidence must be refreshed.

Make invisible authority visible.

Install the free extensions. Run the preflight. Baseline what you have. Add Pro where you need deeper history, authority analysis or runtime mediation.

INSTALL FREE

MCP Trust

Microsoft Marketplace

OPEN ->

INSTALL FREE

Agent Trust

Microsoft Marketplace

OPEN ->

GET PRO

Developer Trust

MCP + Agent Pro

OPEN ->

Resolve before reliance.

ECZ-ID Developer Trust helps you inspect configuration, preserve local evidence, understand authority and re-check current proof. Local policy decides what is sufficient.

developers.ecocitizenz.com/developer-trust

No source, prompt or secret upload for Community checks. MCP Pro mediation is local and user-started. ECZ-ID does not turn configuration evidence into a certification, approval or safety guarantee.