The MCP Security Checklist

Seven questions to ask before you connect customer research to AI tools

Model Context Protocol (MCP) makes it easy to connect AI tools to the rest of your stack. With MCP, your stakeholders can now access insights directly in the AI tools they already use every day. Connecting customer research in this way deserves serious scrutiny. Research data holds participant names, health information, and private opinions people shared with you in confidence.

This checklist helps you protect your participants. If you're considering a customer research platform, ask these questions.

Track answers as you go.
Click below to get an editable version of the MCP security checklist.

Open the checklist

1. Does your MCP inherit our existing permissions, or does it create new access paths?

WHY IT MATTERS:

The biggest security fear is accidental exposure of customer data. If turning on MCP lets anyone with a login see everything in your repository, you could undo years of careful access management.

A GOOD ANSWER SOUNDS LIKE:

"MCP respects your existing role-based access controls. If a user can't see a study in our UI, they can't see it through MCP."

A WORRYING ANSWER SOUNDS LIKE:

"MCP access is workspace-wide" or any pause longer than three seconds.

2. How does authentication work?

WHY IT MATTERS:

A shared API key is a password that nobody owns. It gets copied into files, pasted into Slack, and forgotten on the laptops of former employees.

A GOOD ANSWER SOUNDS LIKE:

"We use per-user OAuth rather than API keys, so there are no long-lived credentials sitting on anyone's device. Revoking someone's platform access revokes their MCP access at the same time."

A WORRYING ANSWER SOUNDS LIKE:

“We will issue your workspace a key.” One credential used by several people means you can’t tell who did what.

3. Will MCP answers surface participant PII data?

WHY IT MATTERS:

Research data is full of personally identifiable information, and you owe it to participants to keep them safe, especially if you’re in healthcare, finance, or government. Once PII travels into an AI tool, it's outside the system that was protecting it.

A GOOD ANSWER SOUNDS LIKE:

No. PII is redacted before any data leaves our server.

A WORRYING ANSWER SOUNDS LIKE:

"Redaction is available if you need it," or any answer implying their detection catches everything.

4. Can AI tools write to our repository, or only read from it?

WHY IT MATTERS:

Read access answers questions. Write access changes your data. Both can be legitimate, but you should know which one you're granting. An AI tool that can edit or delete research data introduces a category of risk worth a much longer conversation.

A GOOD ANSWER SOUNDS LIKE:

"Read and search only. Our MCP tools can't edit, move, or delete anything in your workspace." If write access exists, ask what actions are possible and how the vendor logs them.

A WORRYING ANSWER SOUNDS LIKE:

“It depends on the tool,” or an enthusiastic pitch for agents that update records, with no mention of limits or logging.

5. Is our data used to train AI models?

WHY IT MATTERS:

Your participants didn't consent to becoming training data. Two companies touch your research data when you connect with an MCP — the vendor storing it and the AI provider accessing it. Each has its own policy on training and retention, and a vendor's promise about their own practices doesn't cover the other one.

A GOOD ANSWER SOUNDS LIKE:

"We never train on your data, and that’s written into our contract. It’s worth noting that the AI platform you use has its own separate data and retention policy."

A WORRYING ANSWER SOUNDS LIKE:

“We never train on your data,” full stop. An answer that covers the vendor’s half of the chain and lets you assume it covers the rest.

6. Is there an audit trail?

WHY IT MATTERS:

If research data is ever exposed, you need to see who accessed what and when. Without a log, you can't scope the incident, and you can't prove where it stopped.

A GOOD ANSWER SOUNDS LIKE:

"Every tool call is logged with user attribution, timestamp, and parameters, and your admins can review it directly."

A WORRYING ANSWER SOUNDS LIKE:

“You can request logs from support.” Logs you can’t access on your own are not much use during an incident.

7. What certifications and compliance commitments can you show us?

WHY IT MATTERS:

Certifications are table stakes before you connect your data. They won't tell you whether a vendor designed their MCP server well, but a vendor without them shouldn’t be evaluated at all. Data residency is a hard requirement if you operate in Europe.

A GOOD ANSWER SOUNDS LIKE:

"We're SOC 2 Type II certified, with GDPR and HIPAA commitments in our DPA and EU data residency available. Our current reports and subprocessor list are on our trust page — here's the link."

A WORRYING ANSWER SOUNDS LIKE:

“We are SOC 2 in progress,” or they name certifications in conversation that you can’t verify in written documentation.

The difference between a secure vendor and a risky one

Good MCP security reuses the permissions, authentication, and logging you already trust. If a vendor's answers keep introducing new toggles, keys, exceptions, and governance, the architecture is too risky.

If a vendor answers all seven clearly, congratulations. You've found a platform that treats security as core infrastructure.