PRACTICAL CHECKLIST

MCP Safety Checklist

Before you connect any AI tool to your business or personal systems. For non-technical people connecting AI tools to email, files, calendar, contacts or other systems via MCP, OAuth or any third party integration.

THE STORY BEHIND EVERY ITEM BELOW

A professional in Singapore connected a "verified" Gmail MCP server from a US-based AI marketplace to her email. With one click on an OAuth consent screen, she granted that marketplace full read, write, send, delete and contacts access to her Gmail. She thought she had enabled a useful AI feature. What she had actually done was hand a third party in another country control over her bank statements, passport scans, IRAS notices, CPF letters and password reset emails, because Gmail is the recovery account for almost everything else. She is smart, capable, runs her own business. The problem was not her. The problem was the interface.

BEFORE YOU CONNECT

1. What data will this connector access?

Get specific. Not "your Google account." Which parts? Email body? Attachments? Drive? Calendar? Contacts?

WHY: The Singapore connector could read every email she had ever received. Bank statements, passport photo, every password reset link. She thought it would just organise her inbox.

2. Read only or read-write?

Read only means the AI can look. Read-write means it can also create, modify, send and delete on your behalf.

WHY: The connector she used could send emails as her, delete messages, modify contacts. None of that was needed. Read-write also means an attacker who gets the token can destroy evidence.

3. Who actually built it?

Is the publisher Google, Microsoft, the company that owns the underlying system? Or a third party wrapping someone else's API?

WHY: Her Gmail connector was not built by Google. It was hosted by Smithery, wired into Composio. Three trust boundaries she did not see, all from one click.

4. What permissions is it asking for?

Read the OAuth screen. If it asks for "Allow All" or access to everything, that is a red flag.

WHY: If the tool only needs to read your calendar, it should not ask for full Gmail access. Tools that overreach either do not know what they need or are collecting more than they need.

5. Can you limit the scope?

Some tools let you grant access to specific folders, labels, channels. Choose the narrowest option available.

WHY: If your AI tool only needs to read invoices in one Gmail label, it does not need access to your entire inbox. Most people do not look for the "limit access" option. It usually exists.

6. Does the "verified" badge mean what you think?

Verified usually means the publisher's identity is confirmed. It does not mean the platform itself has been audited.

WHY: In June 2025, the marketplace in this story had a security incident exposing 3,000+ hosted MCP apps. In November 2025, four further CVEs disclosed including account takeover paths. The badge would not have warned her.

7. Where will your token be stored?

Your access token gets stored somewhere. Your laptop? A third party database in another country? Find out.

WHY: Her token now sits in a US marketplace's database alongside thousands of others. If it gets breached, every token is exposed at once. Tokens on your device only put you at risk. Tokens in a shared database put everyone at risk together.

8. Who in your organisation approved this?

If you are connecting tools to a business account, someone should formally approve. If no one has, you have a governance gap.

WHY: This is exactly how the Vercel breach happened. One employee clicked "Allow All" on an AI tool's OAuth. Nobody else knew. When the AI vendor was compromised, the attacker walked into Vercel.

DATA AND PRIVACY

9. Where does the data go?

Which country is it stored in? Which provider processes it? Does it leave Singapore (or your jurisdiction)?

WHY: Singapore's PDPA places obligations on you for how data is handled, even when a third party processes it. If your customer data ends up in a country with weaker protections, you may be breaching your own privacy commitments.

10. Does the vendor train models on your data?

Read the terms. If yes, your confidential information may end up in the model's outputs for other users.

WHY: You may be paying to use a tool while also providing it free training data. If a customer's data appears in someone else's AI output, that is a PDPA breach you caused.

11. Have you tested it with non sensitive data first?

Before connecting real customer information, run the AI with made up data. Watch what it does.

WHY: You will see what it accesses, what it sends, what it returns. Five minutes of testing with fake data prevents finding out the hard way that the AI is doing something you did not intend.

12. Can you revoke access easily?

Can you disconnect within minutes? Test this before you need it.

WHY: When a vendor gets breached, every minute the connection stays live is another minute the attacker has access. Knowing how to revoke is half the battle.

ONGOING GOVERNANCE

13. Is there logging of what the AI accessed?

Can you see an audit trail of what was read, modified, sent or deleted? If no logging, you are flying blind.

WHY: If a connector gets misused, the audit trail is the only way to know what happened. Some MCP servers provide logging. Many do not. Assume invisible if not.

14. Is there a review date set?

Set a calendar reminder to reassess every connection quarterly. Permissions change. Vendor terms change.

WHY: The connection set up in March may have new permissions added in June without you noticing. Quarterly review catches drift before it becomes an incident.

15. Does cyber insurance cover AI tool breaches?

Most policies have exclusions for unauthorised software. If your team connected AI tools without IT approval, check coverage.

WHY: If an AI vendor compromise leads to a breach on your end, your insurer will ask whether the connection was approved and documented. Get this in writing now, not after the incident.

16. Do you have an incident response plan?

If the AI vendor gets breached, what is your first move? Who do you call? How fast can you revoke?

WHY: The Vercel team had to scramble when Context.ai was compromised. The Singapore professional in this story would have to do the same if Smithery is breached. A 30 minute incident plan written today saves you 30 hours of panic later.

YOUR SCORE

Score: ___ / 16 Date: _____

0 to 6: High risk. Stop and review every connection you have made. **7 to 11:** Gaps to close. Prioritise the biggest exposure first. **12 to 16:** Strong position. Keep reviewing.

WHAT TO DO RIGHT NOW

If you have ever clicked "Allow" on an AI tool's OAuth request without reading it carefully, take ten minutes today to:

- → Go to your Google account → Security → Third party apps with account access
- → Look at the list. Anything you do not recognise, or do not actively use, revoke
- → For everything you keep, check what permissions it has
- → Do the same for Microsoft 365, Slack, GitHub, Dropbox, any platform you use

WHEN TO ASK FOR HELP

AI Connector Audit
Independent review of every OAuth grant and MCP connection in your organisation. Find what is exposed.

Vendor Risk Review
Assess the AI marketplaces and third party platforms holding your tokens. Identify trust chain risks.

Incident Readiness
Tabletop exercises for AI tool compromise scenarios. Playbooks for revocation under pressure.

Governance Setup
PDPA aligned governance for AI tool adoption. Approval workflows, registers and quarterly reviews.

TALK TO US

If you are unsure where to start, get in touch. The first conversation is free.
connect@dtnadvisory.com · dtnadvisory.com

Further reading on dtnadvisory.substack.com: "The Safety Briefing No One Gave You" · "Knock Knock, Your AI Tool Just OAuth'd the Attacker In" · "If You Are Having a Bad Day, Read This".