

PRACTICAL GUIDE · COMPANION DEEP DIVE

AI Security, In Depth

A practical guide created by the International Superyacht Society

CURRENT AS OF MAY 2026

SUPERYACHTSOCIETY.ORG

Disclaimer: This is a practical guide, not legal advice, and does not replace your firm's compliance obligations.

If the one-pager is the what, this is the why. AI has quickly become one of the most valuable tools in our work, and the people who use it best are not the cautious holdouts; they are the ones who understand what is happening under the hood and build good habits early. That understanding is what turns AI from a risk into an advantage.

This companion guide goes a layer deeper for two groups:

- 1** The person who has used a chatbot a handful of times and wants to understand the ground rules, and
- 2** The person already building with these tools every day who has been moving fast without looking down.

PART 1

The Five Things to Lock In First

If you have only used AI a handful of times, these five ideas are the foundation.

PART 1

THE FIVE THINGS TO LOCK IN FIRST

If you have only used AI a handful of times, these five ideas are the foundation.



1

PUBLIC MEANS PUBLIC

Treat everything you type as potentially public. Don't put anything in a free tool you wouldn't say out loud in a public place.



2

MARKETING ≠ LEGAL TERMS

The homepage promise is not the contract. Real protection comes from a Data Processing Agreement on enterprise plans—not on Pro, Plus, or Team.



3

AI CAN BE CONFIDENTLY WRONG

AI predicts the next words, it doesn't verify facts. Use it as a first-draft and thinking partner, never as the final source. Always verify.



4

KNOW WHAT AI CAN SEE

Browser tools are limited to the tab. Desktop apps and local agents can access your files—understand and control what you grant.



5

BUILD GOOD HABITS EARLY

- Turn off training-data sharing wherever the option exists.
- Do not connect AI to your email until you understand the scope.
- Use a dedicated browser profile for AI tools and extensions.



THE BIG IDEA

Be deliberate about what goes in, understand what's at stake, and verify what comes out.

1 Treat everything you type as potentially public

On free tiers, assume that what you type can be accessed, whether for quality checks, training, or review. That is not paranoia; it is the agreement you accepted. The simplest discipline for a beginner is to put nothing into a free tool that you would not say out loud in a public place. This is not a reason to fear the technology. It is a reason to be deliberate about what goes in.

2 The marketing line and the legal terms are different documents

“We never train on your data” on a homepage is not the same as the data-processing language in the contract. Anonymized data, embeddings, and AI-generated summaries of your inputs can all be fair game under the terms even when “your data” technically is not. The only real contractual protection is a Data Processing Agreement, and that lives on enterprise plans, not on Pro, Plus, or Team. Those paid tiers buy compute and features; they do not buy contractual data protection. Free tier means public tool. If you need genuine protection, that is an enterprise conversation.

3 Hallucination is structural, not a glitch

An AI model generates the most statistically plausible next words. It does not verify facts and does not know what it does not know. The risk is not that it is always wrong; it is that it is confidently wrong. Use it as a first-draft partner and a thinking partner, never as the final source. Verify anything factual before you publish it, send it, or act on it.

4 Know what AI can actually see on your machine

Browser-based tools are contained to the tab in front of you. Desktop apps and local agents have file-system access, and what they can reach is whatever you approved at install. If you are new to this, do not run a local agent yet. Start in the browser, get fluent, and understand exactly what you are granting before you move to anything that lives on your hard drive. When people ask whether having an AI agent on their computer is too much access, the honest answer is that it depends on what you granted, and most people have never audited that.

5 The one habit to set before you open AI tomorrow

If you take away a single thing:

- ◆ Open the settings on every AI tool you use and turn off training-data sharing wherever the option exists.
- ◆ Do not connect AI to your email until you understand the scope you are about to grant.
- ◆ Use a dedicated browser profile for AI tools and extensions.

That is the floor. Everything else builds on top of it.

PART 2

For the Daily User Already Building

For the people already in the tools every day, here are the areas where speed tends to outrun caution.

PART 2

5 ESSENTIAL CHECKS FOR DAILY AI USERS



1 CONNECTED-TOOL EXPOSURE

Audit connected apps—especially Gmail.
Review what has access, then scope it down or revoke.



2 AGENTS ON YOUR MACHINE

Check which folders your agent can reach.
Move sensitive files out of reach.
Use a dedicated or sandboxed environment for agent work.



3 PAID PLAN ≠ PROTECTED PLAN

Pro, Plus, Team are not protected.
Data protection (DPA) comes with enterprise plans. Read the actual terms.



4 PLUGINS & SUPPLY CHAIN RISK

Vet before you install.
Know the source. Review docs.
If in doubt, don't install.



5 KEYS & CREDENTIALS: PROTECT THEM

Never hardcode keys. Use a secrets manager.
Prefer OAuth & MCP. Watch for prompt injection.
Scope to least privilege and rotate keys.
Protect your customers and your reputation.



THE BIG IDEA

*Move with speed, not at the cost of your data.
Audit, minimize, and protect—every time.*

1 Connected-tool exposure

Gmail integration is the clearest example. When you click “allow” on AI-to-Gmail, you did not grant access to one message. You granted access to the whole inbox: every contract negotiation, every personal exchange, every password-reset email ever sent. Audit it today:

- 1 Go to your Google Account, then Security, then third-party apps with account access.
- 2 Look at what is actually there.
- 3 Scope it down or revoke it.

2 Agents on your primary machine

When you install an agent, the permissions dialog is the moment that matters, and most people click through on the default, which usually grants broader access than they would allow if they read it. The fixes:

- ◆ Audit which directories your agent can reach right now.
- ◆ Move sensitive material (tax returns, medical records, client files, contracts) out of agent-reachable folders.
- ◆ Longer term, use a dedicated machine, a separate user account, or a sandboxed environment for agent work. A clean device with no real data on it is a sound way to experiment.

3 A paid plan is not a protected plan

This is the misconception that catches the most people. Pro, Plus, and Team buy compute and features. They do not include a Data Processing Agreement (DPA). Contractual data protection lives on enterprise plans with a signed DPA. Read your actual data terms, not the sales page. If you are running anything sensitive (health, legal, financial, regulated-industry, or EU-citizen data) through a non-enterprise tier, you are operating outside the protection you think you have.

4 Open-source plugins and supply-chain risk

Community plugins, agent marketplaces, and third-party extensions have no central vetting, and the threat profile is malicious code running with full agent permissions. Vet before you install:

- ◆ **Verified source:** who published this?
- ◆ **Documentation:** is it documented, or is it a black box?
- ◆ **Audit trail:** can you trace the code?

If you cannot tell who built it or why, do not install it.

5 API keys, credentials, and the builder's protocol

For everyone shipping something:

- ◆ Never hardcode keys in your code, ever.
- ◆ Use a secrets manager. Prefer OAuth and MCP wherever the option exists.
- ◆ .env files beat hardcoding, but they are still exposed if pushed to a repo or left on a shared machine.
- ◆ Watch for prompt injection: agent tools can be tricked into exposing keys or taking actions. Scope to least privilege and rotate keys.
- ◆ In customer-facing builds you are protecting both sides of the transaction. Handle customer data carefully, manage secrets properly, and disclose what AI you used. This is your reputational floor.

Two Principles That Tie It Together

Data protection is part of compliance, not a separate task.

Balancing privacy obligations such as General Data Protection Regulation (GDPR) against Know Your Customer (KYC) requirements is core work, not an afterthought. Any AI tool that touches client data falls inside that obligation.

Human judgment stays at the center of regulated work.

Even as the industry adopts AI across KYC and Anti-Money Laundering (AML), the model is that AI supports the experts who handle validation, the difficult decisions, and the exceptions. AI speeds the work; it does not own the decision. That is how all of us should use it.



— BACKGROUND —

WHAT A DATA PROCESSING AGREEMENT ACTUALLY IS



A Data Processing Agreement (DPA) is a formal contract that sets out how a service provider handles, processes, and protects personal data on behalf of another organization.

It draws a clear line between:

**DATA CONTROLLER**

Decides why and how data is processed.

**DATA PROCESSOR**

Handles data on the controller's instructions.

Under regimes like GDPR, having one is frequently a legal requirement, not a nice-to-have.

A SOLID DPA GENERALLY COVERS:**1. SUBJECT MATTER AND DURATION**

What data is processed and for how long.

**2. DATA TYPES AND CATEGORIES OF PEOPLE**

From basic contact details to sensitive financial or health records.

**3. RIGHTS AND OBLIGATIONS**

The rights and obligations of each party, including security and confidentiality duties.

**4. BREACH PROCEDURES**

Breach procedures, with notification typically required within **24 to 72 hours** of discovery.

**5. AUDIT RIGHTS**

Audit rights, so the controller can verify the processor is doing what was agreed.

**THE
TAKEAWAY**

If you are putting personal or client data into an AI tool, the question of whether a DPA exists, and what it says, is not a technicality.

It is the difference between a contractual safeguard and an assumption.



KEY TERMS



1. Data Processing Agreement (DPA):

A legal contract with an AI vendor specifying how your data is handled, stored, and protected. Lives on enterprise plans; Pro, Plus, and Team tiers do not include one.



2. Hallucination:

When a model produces information that sounds plausible but is factually wrong. Not a glitch; it is how the model works, predicting likely next words rather than verifying truth.



3. File-system access:

The permission an installed app or agent has to read, write, or change files on your computer. Browser tools do not have this; desktop apps and agents do, at whatever scope you granted.



4. Supply-chain attack:

When malicious code reaches you through a third-party tool, plugin, or dependency you installed. Especially relevant for unvetted community plugins and agent marketplaces.



5. Prompt injection:

When an attacker hides instructions in a document, email, or webpage to trick an AI into taking unintended actions or leaking information.



6. API key:

A credential that authenticates your code with an outside service. If exposed, anyone can use that service on your account, your bill, and your permissions.



7. Secrets manager:

A dedicated tool for storing keys, passwords, and credentials securely, keeping them out of your code and .env files.



8. Self-hosted model:

Running a model on your own hardware (for example Ollama, vLLM, or LM Studio) instead of calling a cloud API. Maximum data control, with a real productivity trade-off.



9. Zero-retention API:

An endpoint configured so the vendor does not keep your inputs after returning a response. Read carefully; the technical guarantee and the marketing claim are not always the same.



10. OAuth / MCP:

Authentication standards that let you connect tools without sharing raw credentials. Preferable to embedding keys directly in code.



USE THESE TERMS. PROTECT YOUR DATA. STAY IN CONTROL.

Understand the language so you can make better decisions, ask sharper questions, and build with confidence.



THE BOTTOM LINE

DEEPER KNOWLEDGE. SMARTER USE. GREATER ADVANTAGE.



The depth here is not meant to slow you down.

It is meant to make you the kind of user who moves fast, because you understand what is happening underneath.



EVERY SAFEGUARD IN THIS GUIDE EXISTS SO YOU CAN SAY YES TO AI MORE OFTEN, not less—knowing exactly what you have set up and what you are putting at stake.



1. START WHERE YOU ARE

Begin with the basics that fit your needs and your risk.



2. BUILD HABITS THAT FIT

Make security and good judgment part of your workflow.



3. REVISIT AS TOOLS CHANGE

The tools will evolve. So should your safeguards.



Used with that kind of awareness, AI is not a liability to manage.

It is one of the sharpest advantages available to a careful professional.



STAY INFORMED. STAY CURIOUS. STAY IN CONTROL.

The better informed you are, the faster and safer you can move with confidence.



DISCLAIMER

This is a practical guide, not legal advice, and does not replace your firm's compliance obligations. Any views or opinions expressed are solely those of the author and do not necessarily represent the views of the International Superyacht Society or any of its employees, officers or directors.

Images in this presentation were generated using AI.

SOURCES

McKinsey & Company, How agentic AI can change the way banks fight financial crime (2025); Southern Sky AI, AI Security Landmark for Maritime (2026) and AI Tools Data Policy Guide v1.0 (2026); LexisNexis, Know Your Customer (KYC) Explained; Plaid, What KYC is and why it matters in financial services (2025); Icertis, What Is a Data Processing Agreement (2025); H. Nissenbaum, Privacy as Contextual Integrity (2004).

International Superyacht Society · Promoting excellence in the superyacht industry