A GUIDE FOR FINANCIAL ADVISERS · MAY 2026

The AI Model Guide for Financial Advisers.

A comprehensive comparison of the leading AI models on real Australian financial planning scenarios. Covering the results, each model's strengths and weaknesses, how to set up each platform for financial advice work, and ready-made prompts for common adviser tasks.

MODELS TESTED

Claude ChatGPT Gemini M365 Copilot

FOCUS

Australian Financial Advice · May 2026

MODELS

Claude · ChatGPT · Gemini · Copilot

INSIDE

What's in this guide.

SECTION 01

General models aren't purpose-built for financial advice.

If you tried AI tools two or three years ago, you wouldn't recognise them today. The improvement has been that fast. The baseline problem is mostly solved. Ask any major model for the concessional contribution cap or how the assets test works, and it'll generally get it correct.

The real differentiators now are whether a model can handle scenarios where multiple rules interact, whether it uses current-year figures or ones from two years ago, whether it flags the edge case that changes the advice, and whether it pushes back when something doesn't add up or just gives you the textbook answer and moves on.

This guide tests the leading AI models on real-world financial planning scenarios to find out which gives the most useful, accurate, and practice-ready output. It then walks through how to set each platform up for success.

WHY THIS MATTERS

The wrong AI output doesn't just waste time. In this study, models gave answers that differed by tens of thousands of dollars on the same client scenario.

WHO THIS GUIDE IS FOR

Financial advisers, paraplanners and client service officers. Beginner-friendly, with enough depth for experienced users.

WHAT THIS GUIDE ISN'T

This isn't a recommendation to use AI for advice. It's a practical look at what the tools can do today, where they fail, and how to get better results if you're already using them.

1.1 · METHODOLOGY

Same questions. Five models.

THE MODELS

Claude	ChatGPT	Gemini	Copilot
Claude. Sonnet 4.6 Tested with and without skill files.	**ChatGPT. GPT-5.3** No Custom Instructions or Memory.	**Gemini 3 Thinking** Web app with Thinking Mode enabled.	**Copilot. M365 Premium** Auto-thinking enabled, default model.

SCORING 0 TO 5

5 Fully correct · current figures · edge cases flagged · SOA-ready.

4 Correct with good detail, minor gaps only.

3 Core answer right but vague, generic, or missing important nuance.

2 Partially correct but includes a meaningful error or misleading omission.

1 Mostly wrong, outdated figures, or fundamentally misunderstands the rule.

0 Completely wrong or fabricated answer.

TOPIC MIX

Scenarios were designed to cover every common area of financial advice an Australian adviser encounters. Each scenario carried the same weighting regardless of topic area. More complex areas carried more scenarios to properly test the depth of each model's reasoning.

CONTROLS

Same questions for every model. One scenario at a time in a single chat. Fresh chat with no prior conversation history. No memory or custom instructions enabled, except Claude with Skills which had the FP skill files loaded. No prompt template beyond the scenario itself.

VERIFICATION

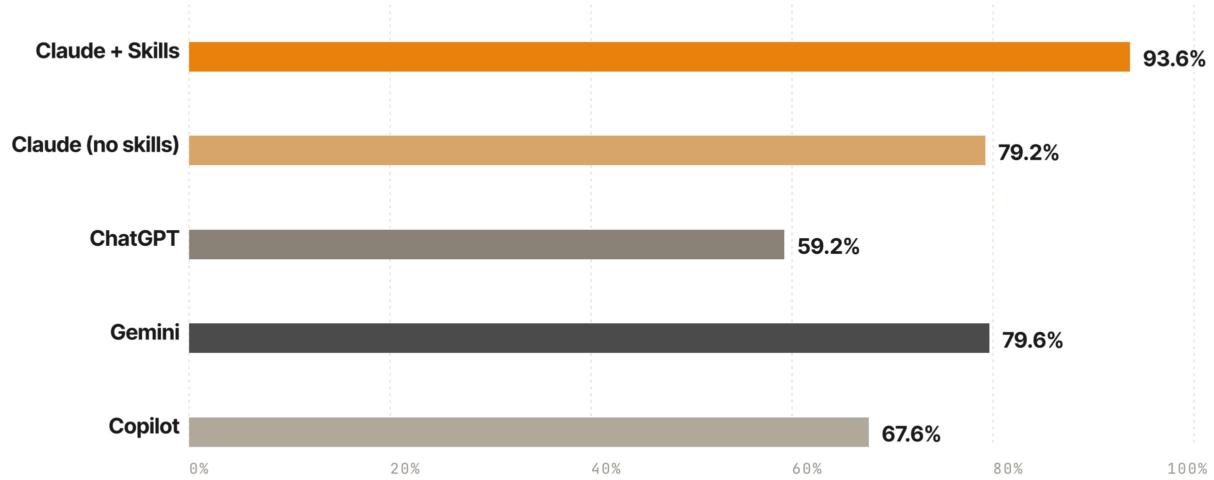
All responses were scored against current legislation, ATO guidance, and verified industry publications as the primary references. Where a model outperformed expectations or a score was borderline, the response was cross-referenced against a second source before finalising.

1.2 · OVERALL RESULTS

A 34% variance on the same questions.

Every model handled the basics. The separation came from how they handled the real work. Using current figures instead of last year's, giving a recommendation instead of just listing options, flagging the interaction you didn't ask about. Following through on the detail that changes the advice.

How each model scored



WHY GEMINI AND CLAUDE TIED

Nearly identical scores, but the responses looked nothing alike. Gemini's strength was strategic thinking. It would often recommend a better overall approach rather than just answering what was asked. Claude without skills was stronger on working through multi-step calculations and applying rules precisely. Each model won different topic areas.

WHAT THE SPREAD MEANS

The gap between 93.6% and 59.2% is the difference between an answer you can work with and one that needs to be rewritten. At the bottom end, models were getting tax calculations wrong, missing concessions entirely, and applying charges that were abolished years ago.

1.3 · BY TOPIC

Every model has topics it handles well.

Claude with skills was the only model with access to verified, current reference material. Every other model relied on its training data and web searches, where the quality of the source was a coin flip. An outdated article or a stale training set meant outdated figures, regardless of how well the model could reason.

That makes the top-line result unsurprising. Give any model better source material and its output improves. But even without skill files, Claude and Gemini both performed strongly, trading wins across different topic areas. The skill files didn't create capability that wasn't there. They made the output consistently accurate across every topic, rather than accurate in some and unreliable in others.

Score by topic area (%)

	Claude + Skills	Claude (no skills)	ChatGPT	Gemini	Copilot
Super contributions	83%	80%	47%	83%	53%
Pensions	95%	82%	57%	80%	68%
Death benefits	100%	60%	53%	60%	70%
SMSF	95%	80%	70%	85%	80%
Tax & structures	90%	80%	70%	85%	75%
CGT	95%	100%	45%	80%	55%
Centrelink	100%	75%	55%	85%	65%
Insurance	90%	85%	70%	80%	75%
FBT	95%	80%	60%	70%	70%
Estate planning	95%	75%	70%	90%	70%
Aged care	90%	80%	70%	90%	70%

95–100 · 80–90 · 60–75 · 50–60 · Below 50

STANDOUT EXAMPLE · 01

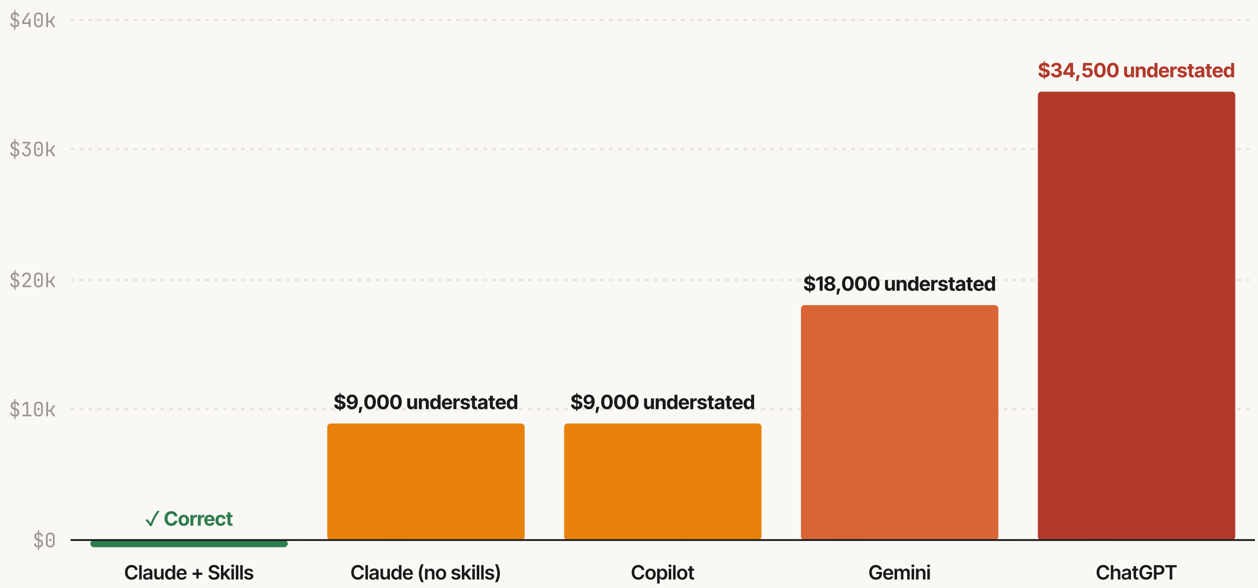
A $34,500 difference in estate planning advice.

THE SCENARIO

Client died at 45 with $800,000 in super. Fund claimed deductions on life insurance premiums. Adult daughter is sole beneficiary. Calculate the taxed / untaxed elements and the tax she'll pay.

Only Claude with Skills correctly applied the legislative formula. The rest either applied it incorrectly, swapped the elements, or didn't attempt the formula at all. The largest error understated the daughter's tax bill by $34,500, a gap that matters if it makes it into a file note unchecked.

Daughter's tax liability (Error by model)



CORRECT TAX

$175,600

22% effective rate on total benefit. Worked from s307-290 service-days formula.

ADVISER TAKEAWAY

If a model skips the formula and estimates instead, the output can be off by tens of thousands of dollars. Complex calculations are where the gap between models is widest.

1.4 · SKILLS IMPACT

Skills give Claude precision. They don't upgrade its reasoning.

Skills are reference documents that Claude reads before answering. current figures, legislative formulas, decision frameworks. They fix **precision**. They don't fix **reasoning**. That's why the improvement isn't evenly spread across topics.

CGT was the one topic where skills actually reduced the score. Without them, Claude had to reason from first principles and produced a more thorough explanation. With them, it leaned on the reference material and skipped some of the working.

NO SKILLS

79.2% →

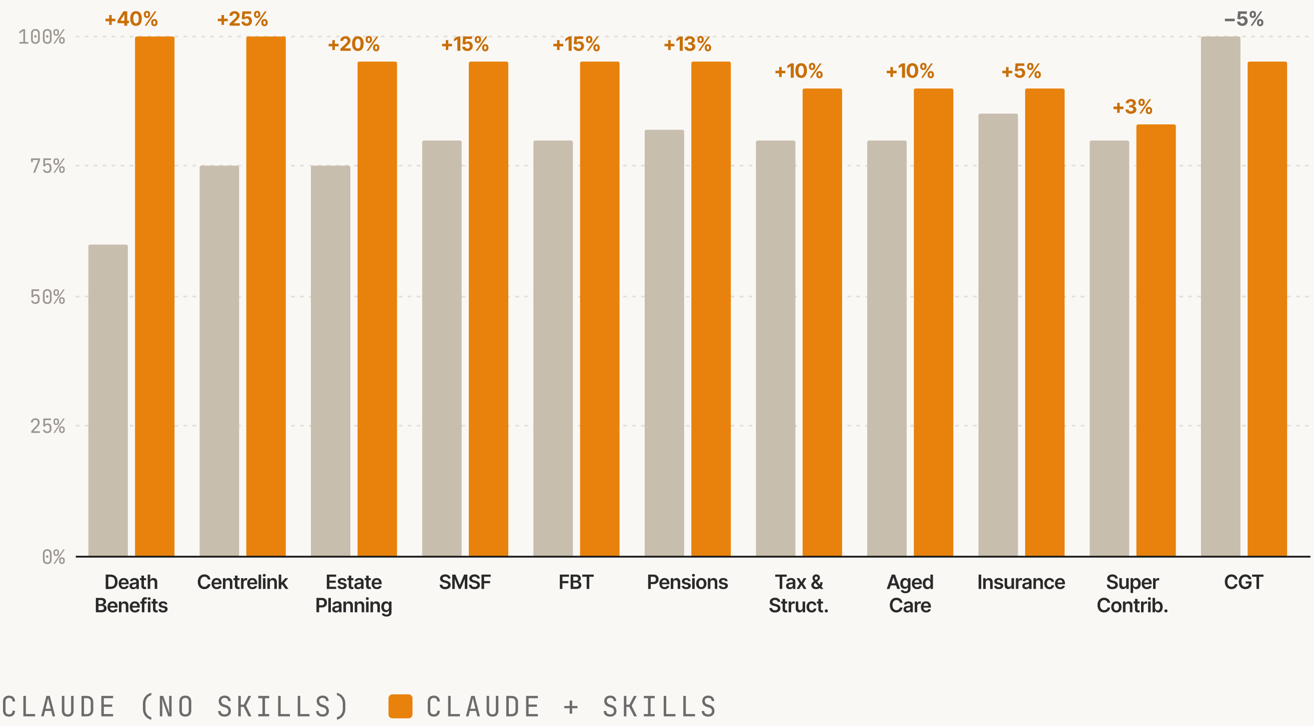
WITH SKILLS

93.6%

+14.4 POINTS

Even with Skills, improvements weren't even across topics. they concentrate where precision and current figures matter most.

Claude with Skills vs without by topic area



SECTION 02

Model strengths & weaknesses.

A profile of each model. What it handles well, where it falls short, and how the scores break down by topic area.

Claude
Claude
93.6% · 79.2%

ChatGPT
ChatGPT
59.2%

Gemini
Gemini
79.6%

Copilot
Copilot
67.6%

MODEL PROFILE · 01

Claude Sonnet 4.6

Claude

POSITIONING

Currently regarded as the strongest model for long-form document analysis, coding, and structured technical work. Widely used by professional services firms, developers, and technical writers. Pro plan is $20 USD/month.

KEY PLATFORM FEATURE

Skill files, reference documents uploaded to Claude that it reads before answering. This test used 16 skill files covering every core area of Australian financial planning.

WITH SKILLS

93.6%

Highest score in the test.

WITHOUT SKILLS

79.2%

Same model, no customisation.

↗ STRENGTHS (IN THIS STUDY)

- Only model to correctly apply complex legislative formulas. like the s307-290 service-days calculation for death benefit untaxed elements.
- With Skills, used correct current-year figures throughout. deeming rates, NCC thresholds, transfer balance cap.
- Without Skills, strong first-principles reasoning. scored 100% on CGT from training data alone.
- Consistently flagged edge cases other models missed.
- Responses structured like genuine file notes rather than textbook answers.

↘ WEAKNESSES (IN THIS STUDY)

- Without Skills, used prior-year figures. deeming rates.
- Without Skills, missed legislative formulas requiring precise application (death benefit calcs).
- Can over-explain when a shorter answer would do.
- Skills took significant effort to build. not plug-and-play.
- Occasionally got historical figures wrong even when current ones were correct.

MODEL PROFILE · 02

ChatGPT GPT-5.3

ChatGPT

POSITIONING

The most widely adopted AI model globally, with the largest user base and third-party integration ecosystem. Commonly the first AI tool people try. Free tier available, Plus is $20 USD/month. OpenAI has been expanding into enterprise and agent capabilities.

KEY PLATFORM FEATURES

Custom GPTs for tailored assistants, **Memory** that retains details across conversations, and the widest third-party integration ecosystem of any model.

↗ STRENGTHS (IN THIS STUDY)

- Clean, well-structured output. consistently easy to read and logically organised.
- Adequate on well-established concepts like insurance definitions, debt recycling mechanics, and basic estate planning frameworks.
- Accessible language that could be adapted for client-facing communication with minimal editing.

↘ WEAKNESSES (IN THIS STUDY)

- Multiple critical factual errors. stated a client with 22 years of ownership didn't meet the 15-year exemption threshold.
- Referenced charges and rules that have been abolished, including the excess concessional contribution charge (removed 1 July 2021).
- Incorrectly stated that TTR pension payments are taxed at marginal rates for a client over 60. they're tax-free from a taxed fund.
- Didn't attempt complex legislative formulas, using "rules of thumb" instead. produced the largest dollar error of any model on the death benefit calculation.
- Outdated figures throughout. used prior-year contribution thresholds and caps.

MODEL PROFILE · 03

Gemini 3 Thinking

Gemini

POSITIONING

Google's latest model, integrated deeply with Google Workspace. Gmail, Docs, Sheets, Drive. Natural choice for organisations already in the Google ecosystem. Google One AI Premium is $32.99 AUD/month.

KEY PLATFORM FEATURES

Gems (custom personas), **Deep Research** mode, and **Thinking Mode** for step-by-step reasoning. Workspace integration for practices already on Google.

↗ STRENGTHS (IN THIS STUDY)

- Best strategic thinker of any model tested. often recommended a better overall strategy rather than just answering the technical question asked.
- Good current awareness. referenced recent case law and legislative changes other models missed.
- Consistently asked useful follow-up questions that an adviser would actually want answered.
- Matched or beat the top score on estate planning (90%) and aged care (90%).

↘ WEAKNESSES (IN THIS STUDY)

- Struggled with complex tax calculations. correctly identified the right approach but made labelling or sequencing errors that changed the dollar outcome.
- CGT reasoning sometimes confused, arriving at the right answer through unnecessary steps.
- Missed FBT nuances around when salary packaging actually works against the client.
- Occasionally presented speculative claims about future legislative changes as fact.

MODEL PROFILE · 04

Copilot M365 Premium

M365 Copilot

POSITIONING

Microsoft's AI assistant embedded across the M365 suite. Word, Excel, Outlook, Teams. The enterprise-first option for organisations already running Microsoft. Primary adoption is through existing M365 Business Premium subscriptions or as an add-on.

KEY PLATFORM FEATURES

Notebooks, a Memory layer for retaining context across sessions. M365 integration means if your practice already runs on Microsoft, Copilot works inside your existing documents and workflows.

↗ STRENGTHS (IN THIS STUDY)

- Good compliance framing. regularly referenced specific legislation sections and linked to source material.
- Practical, adviser-friendly tone that sat well between technical and client-facing.
- Strong on foundational rules across most areas. SMSF (80%), Tax & Structures (75%), Insurance (75%).
- Source linking is genuinely useful for file notes and compliance documentation.

↘ WEAKNESSES (IN THIS STUDY)

- Used outdated NCC bring-forward thresholds, underestimating a client's contribution capacity by $120,000.
- Missed the 15-year small business CGT exemption despite clear eligibility. went through inferior concession stacking instead.
- Less detailed strategic analysis than Claude or Gemini on complex multi-step scenarios.
- Misinterpreted cap levels on the excess contribution cascade question, leading to an incorrect breach conclusion.

SECTION 03

Setting up your platform.

How to set up Claude, ChatGPT, Gemini, and Copilot so you get the most out of the model. plan recommendations, step-by-step setup, and the Beginner mistakes to avoid.

Claude

3.1
Claude
Projects · Skills

ChatGPT

3.2
ChatGPT
Custom GPTs · Memory

Gemini

3.3
Gemini
Gems · Thinking Mode

Copilot

3.4
Copilot
Notebooks · M365

3.1 · SETUP

Setting up Claude.

Claude

HOW THE FEATURES STACK UP

Memory knows who you are. It's passive. Builds up across all your conversations over time ("Australian financial adviser, prefers specific wording"). Available on all plans.

Projects are dedicated workspaces for a type of work. Each has its own instructions and uploaded files. Every chat inside reads them automatically. Think of it as briefing a specialist for a particular job. Available on all plans, though Pro gives you higher usage and access to the better model.

Project knowledge is the reference material sitting inside a Project. SOA templates, past examples, firm-specific style guides. The playbook for this specific kind of work.
Skills are portable reference files that load dynamically when Claude needs them, and work everywhere, not just inside one Project. FY figures, legislation, decision frameworks. The textbooks on the specialist's desk, available across every job. Uploaded once at Customise → Skills. Available on all plans.

STEP-BY-STEP

1. Sign up at **claude.ai**. Pro (A$34/month) unlocks Opus, higher usage, and priority access during peak times.
2. Click **Projects** in the sidebar → create one (e.g. "Financial Advice").
3. Add Project instructions. tone, structure, financial year, things to always flag.
4. Upload the FP reference files (linked at the back) plus your own examples and templates.
5. Start every new conversation *inside* the Project.

WHY THE COMBINATION MATTERS

You don't need all three. Memory alone helps. But the stack is what produced the 93.6% result in this test. Memory knows who you are. The Project knows what job you're doing. The Project knowledge gives it the technical detail to do it properly.

BEGINNER MISTAKES

- Starting chats outside the Project, no instructions, no files, generic output.
- Uploading files without Project instructions, so Claude doesn't know how to use them.
- Stacking unrelated scenarios into one chat.
- Asking vague questions without the relevant details.

3.2 · SETUP

Setting up ChatGPT.

ChatGPT

HOW THE FEATURES STACK UP

Custom Instructions, set manually in Settings → Personalisation. Two text boxes: what ChatGPT should know about you, and how it should respond. Applies to every conversation. Your standing brief. Available on all plans.

Memory, builds passively from your conversations, or from things you explicitly tell it to remember. View and edit in Settings → Personalisation → Memory. The colleague who's worked with you for months and just knows how you like things. Free plan gets a lightweight version; Plus and above get full recall.

Custom GPTs, a separate assistant you build, with its own instructions, documents, and personality. Doesn't share Memory or Custom Instructions with main ChatGPT. A specialist, not a generalist, "SOA Drafter," "Client Email Writer." Free users can use existing GPTs from the store; creating your own requires Plus or above.

Projects, persistent workspaces that group related chats with shared files and instructions. Sits between Custom Instructions (too broad) and Custom GPTs (too rigid) for ongoing work. Similar to Claude's Projects. Available on Plus and above.

STEP-BY-STEP

1. Sign up at **chatgpt.com**. Plus (A$35/month inc GST) unlocks Custom GPTs, Projects, and GPT-5 Thinking. Go (A$13/month) is a cheaper tier that includes Projects and Custom GPTs but skips the reasoning models, likely not enough for professional work.
2. Set **Custom Instructions** — "Australian financial adviser. Likes certain wording."
3. Turn on **Memory** so it learns your preferences.
4. Build a **Custom GPT** for a repeatable workflow (Explore GPTs → Create → upload your reference docs).
5. For complex technical questions, switch to the **o-series reasoning model** via the model selector.

MODEL MODES

Default (GPT-5 Instant) is fine for drafting and simple questions. Switch to GPT-5 Thinking for multistep calculations, interacting rules, or anything where you need reasoning rather than retrieval. The Pro tiers above Plus add GPT-5 Pro for the hardest problems, though Plus covers most professional work.

BEGINNER MISTAKES

- Default model on complex technical questions, use Thinking mode.
- Skipping Custom Instructions, so every chat starts from zero.
- Expecting Custom GPTs to inherit your Memory.
- Memory off, then wondering why it never learns.

3.3 · SETUP

Setting up Gemini.

Gemini

HOW THE FEATURES STACK UP

Canvas, interactive workspace alongside chat. Write documents, build apps, create decks, visualise data. Activate via "Canvas" below the prompt bar. Free on Gemini 3 Flash; AI Pro/Ultra get 3.1 Pro with a 1M-token context window. Exports to Slides or PDF.

Gemini Live, voice conversation mode. Talk to Gemini hands-free. Search Live lets you point your camera at something and ask about it in real time.

Google Workspace integration, connects Gemini to Gmail, Calendar, Drive, Docs, and Sheets. Reference existing documents without copy-pasting. If your practice runs on Workspace, this is where Gemini has a genuine edge.

Gems with Docs, custom personas with saved instructions. The key feature: link a Gem to a Google Doc and the Gem's knowledge updates automatically when you edit the source. Point one at your FY figures doc — update it once in July, every conversation after uses the new numbers.

STEP-BY-STEP

1. Sign up at **gemini.google.com**. Upgrade to Google AI Pro (A$32.99/month) for Gems with file references, Deep Research, 1M-token context, and 3.1 Pro. AI Plus (A$12.99/month) is the cheaper mid-tier but caps context at 128k tokens and skips Deep Research, probably not enough for serious document work. AI Ultra (A$409.99/month) sits above for heavy users who want Veo 3.1, Agent Mode, and the upcoming Deep Think.
2. Create a **Gem**, describe it conversationally. Link it to a Google Doc of your FY figures so it auto-updates when you edit the source.
3. Switch to **Thinking (3.1 Pro)** before sending complex questions. The quality lift is significant.
4. Connect **Drive** if your practice is on Workspace. reference Docs directly without uploading.
5. Use **Deep Research** for legislative changes or background for a complex scenario. it runs multiple searches and synthesises a report (AI Pro and Ultra only).

BEGINNER MISTAKES

- Leaving Thinking Mode off on complex questions.
- Assuming Gemini reads your Drive without pointing it to specific files.
- Taking speculative statements about future legislation at face value.
- Setting up a Gem but not linking it to a Doc.
- Expecting AI Plus to do AI Pro work, the context cap and missing Deep Research will bite.

3.4 · SETUP

Setting up Microsoft Copilot.

M365 Copilot

HOW THE MODES WORK

Quick Response gives fast, straightforward answers. Use it for simple lookups where you don't need reasoning.

Think Deeper, up to 10 seconds, uses OpenAI's reasoning models. Use for complex technical questions that need multi-step reasoning.

Smart, runs GPT-5.2 and auto-adapts between quick and deep reasoning. The "let Copilot decide" mode and the sensible default if you're unsure.

Search, prioritises live web results with citations. Use when you need current info or want to verify against sources.

Study and Learn, guides you toward an answer step by step rather than giving it outright. Useful when you're trying to understand a rule, not just apply it.

NOTEBOOKS AND PAGES

Notebooks are persistent AI workspaces. add Word, PowerPoint, Excel, PDFs, even entire SharePoint sites and Copilot reasons across everything. New features include an Overview page, Audio Overviews (podcast-style summaries), Mind Maps, and direct generation of Word docs and PowerPoint decks from notebook content. Copilot's equivalent of Claude's Projects.

Pages are collaborative canvases built on Microsoft Loop. Take chat outputs and turn them into shareable, editable documents. Pages are the output layer; Notebooks are the input and thinking layer.

STEP-BY-STEP

1. Access via your M365 subscription or at **copilot.microsoft.com**.
2. Set up a **Notebook**, add your FY figures, file note templates, and licensee compliance material.
3. In Word, Excel, and Outlook, use the **Copilot sidebar** directly rather than switching to a separate chat.
4. Use **Think Deeper** or **Smart** mode for technical questions; reserve Quick Response for lookups.
5. When the output's ready to share, export to a **Page**.

M365 INTEGRATION

Copilot's genuine differentiator. If your practice runs on Microsoft, it's embedded in what you already use. draft in Outlook, build formulas in Excel, restructure docs in Word, without leaving the app.

BEGINNER MISTAKES

- Quick Response on complex technical questions, use Smart or Think Deeper.
- Copilot in Word for technical questions when the standalone chat performs better.
- Skipping Notebooks and losing context between sessions.
- Assuming M365 integration gives Copilot access to everything. you still point it to specific files.

SECTION 04

Getting better answers from any AI model.

Most of the difference between a useless AI response and a genuinely helpful one comes down to how you ask. The model is the same. your input is the variable.

4.1 · SEVEN HABITS

Seven habits that improve any model.

01 · GIVE IT THE FULL PICTURE

The single biggest mistake is being too vague. "What are the contribution rules for my client?" gives a textbook answer. Compare that to: *"client is 68, retired, TSB $1.4m. She sold an investment property and has $350,000 in after-tax proceeds. How much can she contribute?"*, the second gives the model enough to work with.

02 · TELL IT WHAT YOU WANT BACK

If you want a file note, say so. Dot points for a client meeting? Say that. "Structure your answer as a file note with the client's position, the relevant rules, your recommendation, and any risks to flag" changes output dramatically.

03 · TELL IT WHAT TO CHECK

If you know there's a trap. a Centrelink interaction, a TBC issue, an age-based restriction. mention it. Even "are there any edge cases or traps I should be aware of" is enough to push it beyond the obvious.

04 · USE IT AS A SECOND OPINION

The strongest use case isn't "write SOA." It's "here's thinking. what am I missing?" The models are often better at finding gaps in your reasoning than building the reasoning from scratch.

05 · WHEN IT LOOKS TOO CLEAN, PUSH BACK

A confident answer with no caveats on a complex question is a red flag. Ask "what assumptions are you making?" or "what would change if the client was a non-resident?" The best models flag complications without being asked. but no model does it every time.

06 · KEEP IT TO ONE SCENARIO

Stacking three client scenarios into one message dilutes the answer quality. One scenario per conversation, especially for anything technical.

07 · IF IT GETS SOMETHING WRONG, TELL IT

The models adjust within a conversation. If it uses the wrong year's figures, say so. If it misses an interaction, point it out. Most of the time it'll correct itself on the second pass.

> BONUS · USE AI TO BUILD BETTER PROMPTS
>
> If you're not sure how to frame a question: *"I need to ask you a question about a client's super contribution strategy. What information do you need from me to give the best possible answer?"* The model comes back with a checklist. Fill in the blanks and send it back.

4.2 · PROMPTS FOR ADVISERS

Prompts for advisers.

Copy, swap in details, adjust to suit. Replace the **orange placeholders** with your client's specifics.

01 · CLIENT SCENARIO ANALYSIS

GENERAL PURPOSE

"client [age, situation] has [balances, assets, income]. They want to [goal]. Walk me through the relevant rules, any interactions between them, and what you'd recommend. Flag any edge cases or traps."

WHEN A client's asked a question, or you're building a strategy and need the technical answer before you start writing anything.

ATTACH Usually nothing. paste the relevant fact-find or meeting-note details. For death benefits or SMSF, having the relevant Skill file in your Project does the heavy lifting.

02 · STRATEGY COMPARISON

DECISION SUPPORT

"client is deciding between [option A] and [option B]. Here's their position: [details]. Compare the two options. tax outcome, Centrelink impact, estate planning implications, and any risks. Which would you recommend and why?"

03 · FILE NOTE FROM MEETING NOTES

ADMIN

"Here are rough notes from a client meeting: [paste notes]. Write this up as a structured file note covering the client's current position, what was discussed, the agreed actions, and any follow-up required."

WHEN You've walked out of a review with a page of scribbled dot points and need to turn it into something your licensee won't send back.

ATTACH Your rough notes. messy is fine. Upload your practice's file note template to the Project once and the model will match the structure every time.

04 · COMPLIANCE CHECK

RISK REVIEW

"Review this strategy for a client: [details and proposed recommendation]. Are there any compliance issues, regulatory requirements I need to document, or risks I haven't addressed?"

4.2 · PROMPTS FOR ADVISERS · CONTINUED

Prompts for advisers.

05 · SOA SECTION DRAFTING — DOCUMENT DRAFTING

"Draft the [strategy section / basis for advice / alternatives considered] for a client with the following position: [details]. The strategy being recommended is [strategy]. Match the tone and structure of this example: [paste example or reference uploaded template]."

WHEN — You know the strategy, you know the client. you just need the words on the page. Best on repetitive sections like basis for advice and alternatives considered.

ATTACH — A previously licensee-approved SOA section, plus the scenario details. The approved example is critical. it's the difference between output you can use and output you have to rewrite.

06 · CLIENT-FACING EXPLANATION — CLIENT COMMS

"I need to explain [concept] to a client who isn't financially literate. Keep it simple, no jargon, and use a practical example."

07 · CHALLENGE THINKING — SECOND OPINION

"Here's draft recommendation for a client: [details]. Play devil's advocate. what am I missing, what could go wrong, and is there a better approach I haven't considered?"

WHEN — You're fairly confident about a recommendation but something feels slightly off, or you've been staring at the same scenario too long and need a fresh perspective.

ATTACH — Your draft recommendation. rough is fine. The more specific you are about what you're recommending and why, the better the pushback. "Check this" gets you generic caveats.

08 · USE AI TO BUILD BETTER PROMPTS — META

"I need to ask you a question about [topic / client's strategy]. What information do you need from me to give the best possible answer?"

A WARNING

99%

THE TRAP

AI models are getting better fast. That's the problem.

When a tool gets most things right, you stop checking everything. The first few times you use it you're reading every line, cross-referencing, questioning the output. Six months in, you're copying and pasting into file notes without a second look.

Even the best-performing model in this test, **scoring 93.6%** — got things wrong. It used the wrong historical CC cap on one question. On another it introduced ambiguity where there shouldn't have been any. These aren't catastrophic errors, but they're the kind of thing that slides into an SOA unnoticed if you're not actively looking.

The more dangerous errors came from other models. stating a client with 22 years of ownership didn't meet a 15-year threshold, referencing a charge abolished years ago, understating a death benefit tax bill by $34,500. These aren't edge cases. They result in wrong advice, client losses, and compliance problems.

THE RULE

No model is a source of truth. Every model is a first draft that needs your professional judgement applied to it. If you build AI into your workflow, build the checking step in too.

NOTES · 1/2

Disclaimers.

LEGAL DISCLAIMER

This guide reflects the author's personal findings and opinions as at May 2026. It is not financial advice, legal advice, or a recommendation to use any specific product or platform. AI models are updated frequently and performance may differ from what is described here.

YOUR LICENSEE, COMPLIANCE TEAM, AND IT DEPARTMENT SET THE STANDARD. NOT THIS GUIDE.

Nothing in this guide overrides your firm's policies on AI use, data handling, or approved software. If your licensee hasn't approved a platform, this guide is not approval. If your compliance team has restrictions on how AI can be used in the advice process, those restrictions apply regardless of what any model scores in a test. Before using any AI tool for client work, confirm it is permitted under your firm's policies. If no policy exists yet, raise it with your compliance team before you start.

AI MODELS ARE NOT RELIABLE SOURCES OF TRUTH.

This is not a limitation that will be fixed in the next update. It is a structural feature of how these models are built. They are trained on large datasets that include outdated, incorrect, and contradictory information. They predict the most likely response, not the most accurate one. Every model tested in this guide produced errors. The highest-scoring model still got things wrong. No model should be used as a substitute for professional judgement, primary sources, or qualified advice.

EVEN THE PLATFORMS THEMSELVES SAY THIS.

Microsoft's terms of service for Copilot describe its outputs as being for entertainment purposes only and state they should not be relied upon as professional advice. This is not a throwaway legal line. Microsoft's own legal team requires that disclaimer because they cannot guarantee the accuracy of what the model produces. The same applies across every platform covered in this guide. If the companies building these models won't stand behind the output, neither should you.

THIS GUIDE IS A TOOL, NOT A RECOMMENDATION.

The purpose of this guide is to help financial advice professionals understand the differences between AI platforms and get better results from whichever one they choose. It does not confirm the accuracy of any model's output. It does not confirm the privacy or security of any platform. It does not guarantee that any platform will handle client data safely. Those assessments are the responsibility of your firm, your licensee, and your IT and compliance teams.

ALWAYS VERIFY. EVERY TIME.

AI is useful when treated as a first draft, a research starting point, or a second opinion. It is dangerous when treated as a finished product. Check the figures. Check the legislation. Check the logic. If you wouldn't sign off on advice without reading it, don't sign off on AI output without checking it either.

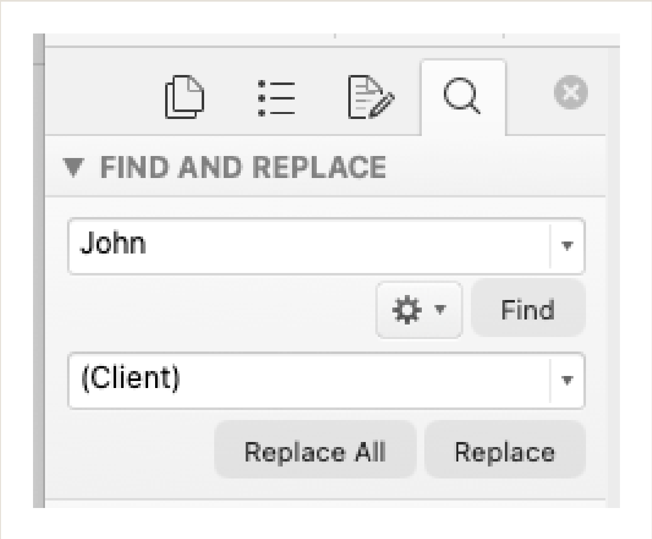
NOTES · 2/2

Privacy.

The AI doesn't need to know who the client is. It needs ages, balances, income, and the scenario. A quick find and replace before you copy anything across is enough.

STRIP OUT

- Client names — replace with "Client 1" or "Client 2"
- Account numbers, member numbers, fund IDs
- Employer names — replace with "Employer"
- Addresses, phone numbers, email addresses
- Dates of birth — replace with just the age
- TFNs, ABNs, any government identifiers



BROADER HABITS

- Don't paste a full SOA or advice document into any model. Extract the scenario details you need and paste those only. The model doesn't need 40 pages of appendices to answer a contribution question.
- Check your licensee's AI policy before using any model for client work. Some licensees have approved specific platforms. Some require disclosure of AI use in advice documents. Some haven't addressed it yet. Find out before you start.
- For sensitive scenarios, check your platform’s data retention settings. Most platforms let you control whether conversations are saved or used for model training.
- Be aware that conversations may be stored by the platform. If your firm has records retention obligations, consider whether AI chat logs form part of the advice record. If you're not sure, ask your compliance team.
- Before pasting, use Find and Replace in Word (Ctrl+H) to swap names, DOBs, and account numbers for generics, “John” → “(Client)”. Keeps identifiable data out of the model.

PLATFORM PRIVACY. QUICK NOTES

Claude

Opt out of training in Settings → Privacy. API isn’t trained on; ZDR available on request.

ChatGPT

Business and Enterprise plans don’t train on your data by default. Individual plans require manual opt-out in Settings → Data Controls.

Gemini

On Workspace, Gemini operates under the same enterprise data policies. Consumer Gemini trains by default; opt out in Activity settings.

Copilot

Microsoft 365 Copilot sits within the M365 security boundary and data policies your organisation already has in place. The free Copilot doesn’t.

FREE DOWNLOAD · THE SKILL FILES

The 16 Skill files that took Claude from 79.2% to 93.6%.

Distilled from about 3,000 pages of verified source material and cross-checked against FY2025/26 technical references. Current-year figures, legislative rules, decision frameworks, edge cases, and common traps.

WHAT A SKILL FILE ACTUALLY IS

A plain-text reference document the AI reads before answering. Each one covers a core area of Australian financial advice, including current-year figures, rules, decision frameworks, and the traps that catch people out.

In this test, they were the difference between a model quoting the correct NCC threshold ($1.76m) and one still on last year's ($1.66m), a $120,000 error.

HOW TO USE THEM

Upload all 16 to a Claude Project once. The model references the right file for whatever scenario you bring it, no swapping per topic, no re-uploading each chat.

They also work in ChatGPT Custom GPTs and Gemini Gems, anywhere you can load reference documents into a persistent workspace. They come with a detailed setup guide to walk you through it.

WHAT'S IN THE SET

Super contributions	Pensions	Tax components	Death benefits
SMSF	Personal tax	CGT	Centrelink
Insurance	Estate planning	Investments	FBT
Ethics	Aged care	Expatriate	FY26 Figures

GRAB THEM HERE

The Skill files are free to download. Find them here or message me on LinkedIn and I'll share the link.

linkedin.com/in/bkingfinance