INFORMATION SECURITY × CLAUDE · A FIELD GUIDE FOR PRACTITIONERS

Claude for InfoSec Teams.

Nine practical workflows — threat research, log analysis, briefings, OSINT, automation — plus five reusable prompt shortcuts, rewritten for how a security team actually works. Open it Monday morning. Run one workflow. Repeat.

THE FIVE SHORTCUTS INSIDE

`/god`	`/ghost`	`/artifacts`	`/blindspot`	`/deepdive`

BUILT FOR SECURITY OPERATIONS, THREAT INTEL, GRC AND ENGINEERING

Share it with your team · Save this

01 · WHY THIS GUIDE EXISTS

AI, the way a security team should use it

Most AI playbooks are written for a general audience. This one is for **security practitioners** — people who handle alerts, advisories, audits, incidents, and executives who want answers yesterday. Every workflow below is aimed squarely at security work, and every prompt is copy-paste ready.

The core mindset shift: stop treating AI like a search engine. Treat it like a tier-1 analyst pool that never sleeps — one that can read a 90-page vendor report, triage a CSV of alerts, draft the incident comms, and build the briefing deck, in parallel. Your job shifts from doing every task to **directing and verifying** the work.

> *“The analysts who stay valuable aren’t the ones who know the most CVEs. They’re the ones who move first and verify fastest.”*

How each module is structured

- **What it is** — the workflow in one line.
- **Security use cases** — where it earns its keep in real security work.
- **The prompt** — copy-paste ready, built on the four-layer context structure.
- **Tools used** — and any data-handling cautions specific to security work.

How to use this document

Don’t read it end-to-end. Pick the one workflow that solves your most annoying problem this week and run it three times. Depth beats breadth — one workflow internalized beats nine skimmed.

NON-NEGOTIABLE GROUND RULE BEFORE ANYTHING ELSE

- Follow your organization’s AI acceptable-use policy. When in doubt, ask before pasting.
- Never paste credentials, API keys, customer PII, or classified/restricted material into any AI tool.
- Sanitize logs and tickets first: strip usernames, internal hostnames, IPs you consider sensitive.
- AI output is a draft, not a verdict. Verify IOCs, CVE details, and config changes against primary sources before acting.

02 · CONTENTS

What’s inside

Nine workflows plus the shortcut layer. Each one independently useful. Together, a faster security team.

01

FOUNDATION

Context Engineering for Security Prompts

The discipline of giving Claude exactly what it needs to nail the output — every time.

The skill isn't "writing better prompts." It's engineering the right **context**. In security work that matters double: a vague prompt gets you generic advice; a structured one gets you something you can put in a ticket, a report, or an exec's inbox.

The four layers

CONTEXT	TASK	INSTRUCTION	DATA
Who Claude is playing — "senior SOC analyst," "GRC lead," "incident commander."	What you want done. One clear sentence.	Format, length, audience, banned content.	Sanitized logs, advisories, policies, examples.

COPY-PASTE PROMPT · INCIDENT NOTIFICATION EMAIL

```
You are a senior information security communications lead who writes clear,
calm, jargon-free notifications for non-technical employees.

Write an all-staff email about a credential-phishing campaign currently
targeting our company, asking employees to report suspicious messages.

Keep it under 180 words. No fear-mongering, no technical jargon, no acronyms
without expansion. Second-person voice. End with exactly two actions: how to
report, and what to do if they already clicked. Do not name any employee.

Context: Attackers are spoofing our HR department with fake "benefits update"
emails linking to a credential-harvesting page. IT has blocked the domain.
Reports go to phishing@company.com or the "Report Phish" Outlook button.
```

THE 5-SECOND CHECKLIST BEFORE YOU HIT ENTER

- Have I told Claude who to be? (Context)
- Is the task one sentence, not a paragraph of wishes?
- Did I specify audience, format, length? (Instruction)
- Did I paste the sanitized source material? (Data)
- Would a smart junior analyst know exactly what to do from this?

Pro tip: if you don't specify the role, you get the average of the internet. "Senior detection engineer at a Fortune 500" is the single highest-leverage line in any security prompt.

02

RESEARCH

Your Threat & Risk Research Intern

Research mode: a consultant-grade analyst for any security topic, in minutes.

Claude's Research mode reads across dozens of sources, cross-references them, and returns a structured, cited report. For security teams, that turns "I'll dig into it next sprint" into "here's the landscape, 20 minutes later."

When to use it

- Threat landscape briefs for a sector, technology, or threat actor group
- Vendor security due diligence before procurement or renewal
- Regulatory scans — what changed in your compliance obligations this quarter
- Tooling evaluations — EDR, SIEM, CSPM market comparisons before a PoC

COPY-PASTE PROMPT · RESEARCH MODE

```
You are a principal threat intelligence analyst who produces structured,
source-cited research using recognized frameworks (MITRE ATT&CK, NIST CSF).

Research the current ransomware threat landscape for mid-size companies in
our industry: active groups, common initial access vectors, dwell times, and
recently exploited CVEs.

Then map the top techniques to MITRE ATT&CK and recommend the five detection
or hardening priorities that would most reduce our exposure.

Cite sources throughout. Flag anything where reporting conflicts. Make the
final section a one-page executive summary a CISO could forward as-is.
```

How to adapt it

- **The role** — swap to "GRC analyst," "cloud security architect," "red team lead."
- **The topic** — your sector, your stack, the vendor under review.
- **The angle** — the decision the research must support (buy, patch, accept, escalate).

TOOL	WHAT FOR	WHERE
Claude — Research Mode	Deep multi-source research with citations	claude.ai → toggle "Research"
Primary sources	Always verify CVEs/IOCs: NVD, vendor advisories, CISA KEV	nvd.nist.gov · cisa.gov

03

OSINT

Social Listening & Open-Source Intel

Find out what attackers, researchers, and victims are saying — in near real time.

Formal reporting lags reality by days. Researcher chatter, vendor blogs, and social posts often surface exploitation in the wild first. The play: gather raw signal with a live-search tool, then have Claude structure it into something your team can act on.

When to use it

- A CVE drops on a product you run — is anyone seeing active exploitation?
- Brand and domain monitoring — typosquats, fake support accounts, leaked-data claims
- Pre-incident sentiment — is your company being named in breach chatter?
- Tracking researcher consensus during a fast-moving event (e.g., a supply-chain compromise)

COPY-PASTE PROMPT · STRUCTURE THE RAW SIGNAL

```
You are a threat intelligence analyst who turns raw open-source chatter into
a structured intelligence brief.

I'm pasting raw posts, headlines, and quotes gathered in the last 24 hours
about [CVE-XXXX-XXXXX / incident name].

Produce: (1) what is confirmed vs. claimed vs. speculation, with your
confidence level for each; (2) indicators or behaviors mentioned, clearly
labeled UNVERIFIED; (3) what this means for an org running [our stack];
(4) three monitoring or mitigation actions for the next 48 hours.

Be skeptical. If sources conflict, say so explicitly.
```

SECURITY CAUTION

- Treat scraped chatter as **untrusted input** — it can contain false IOCs or even text crafted to manipulate AI tools (prompt injection). Ask Claude to label confidence, and verify before blocking or alerting.
- Don’t paste internal incident details into public tools to “compare” — keep the flow one-directional: public signal in, internal analysis stays internal.

TOOL	WHAT FOR	WHERE
Claude + web search	Pull and structure recent reporting with citations	claude.ai
Grok	Live X/Twitter search for researcher chatter	grok.com
Vendor blogs / CERTs	Ground truth to verify the chatter against	cisa.gov · vendor PSIRTs

04

ANALYTICS

Log & Alert Data Analysis

Upload a sanitized export. Get the read that used to take a week of pivot tables.

Drop a CSV or Excel export into Claude — SIEM alerts, phishing-report metrics, vuln-scan results, access reviews — and ask the questions you actually care about. Claude writes and runs the analysis code, charts it, and tells you what stands out.

When to use it

- Monthly alert trends before the metrics meeting — volume, false-positive rate, MTTR
- Vulnerability scan exports — cluster by exploitability, owner, and age, not just severity
- Access review spreadsheets — find dormant accounts, privilege outliers, toxic combinations
- Phishing simulation results — which departments, which lure types, repeat clickers

COPY-PASTE PROMPT · SECURITY DATA ANALYSIS

```
You are a senior security data analyst who turns messy exports into clear
operational insights.

I've uploaded [describe file — e.g. "our Q1 SIEM alert export: 14k rows,
columns for rule name, severity, source, disposition, time-to-close"].

Analyse it and give me:
1. The top 3 findings a CISO would want to know
2. The noisiest rules by false-positive rate — candidates for tuning
3. Anything anomalous: spikes, gaps in logging, impossible values
4. Two concrete recommendations for next quarter
5. One chart that best tells the story; build an interactive dashboard
artifact if the data supports it.

Be direct. Flag data-quality problems instead of working around them silently.
```

BEFORE YOU UPLOAD

- Sanitize first: strip or hash usernames, emails, internal hostnames and IPs unless you've confirmed the data class is approved for the tool.
- Messy files are fine — merged cells, junk headers, totals rows. Mention the quirks; Claude will clean before analysing.

05

CLAUDE SKILLS

Security Briefings in Minutes

Turn research, incident notes, or metrics into a real .pptx or .docx you can present.

Claude Skills generate actual files — PowerPoint, Word, Excel, PDF — not just outlines in chat. Ask for a deck and you get an editable .pptx to download. Chain it with Module 02 or 04: run the analysis, then in the same conversation say “now make the deck.” Claude already has the context.

When to use it

- Quarterly security review decks for leadership
- Incident post-mortems and lessons-learned documents
- Security awareness training slides tailored to a department
- Audit-ready policy documents and control narratives in Word

COPY-PASTE PROMPT · EXEC BRIEFING DECK

```
You are an expert presentation designer who builds board-grade security
briefings: visual, minimal, zero walls of text.

Using the ransomware research and the Q1 alert analysis from this
conversation, create a 10-slide PowerPoint for our executive risk committee.

Cover: threat landscape, our current exposure, what the Q1 data shows,
top 5 priorities, resource ask, and a 90-day roadmap.

Use the pptx skill. Navy blue and white palette, one idea per slide,
plain-English titles that state the takeaway (not "Update" but
"Phishing reports doubled — and that's good news").
```

PRO TIPS

- Specify slide count — “10 slides,” not “a deck.” Claude paces it correctly.
- Specify the aesthetic — “navy, minimal, board-grade.” The default is generic.
- Iterate in-chat: “make slide 4 bolder,” “add a budget slide.” It regenerates the file.
- Same pattern works for Word (.docx), Excel (.xlsx) and PDF deliverables.

06

COMPREHENSION

Summarizing Advisories, Reports & Policies

90 pages of vendor PDF, compressed to the five things you must act on.

The fastest way to 10x your input bandwidth. Drop in a threat report, a SOC 2 report, a pentest deliverable, a new regulation, or a meeting transcript — and pull out exactly what matters in 60 seconds.

When to use it

- Annual threat reports (Verizon DBIR, vendor intel) — extract what applies to *you*
- Pentest and audit reports — findings by real risk, not by the vendor's severity labels
- New regulations or framework updates — what actually changed vs. last version
- Long incident bridge transcripts — decisions made, owners, open actions

COPY-PASTE PROMPT · RUTHLESS EXTRACTION

```
You are a principal security analyst who extracts only what matters from
long documents.

I've uploaded [a pentest report / threat intel report / audit report /
regulation text].

Give me:
1. The single most important takeaway in one sentence
2. The 5 findings most relevant to an org with [our stack/context],
each under 25 words
3. Every concrete number, deadline, or commitment worth remembering
4. The three actions we should take, with a suggested owner role
5. Anything that contradicts what we currently assume or do

Be ruthless. Skip boilerplate, methodology filler, and vendor marketing.
```

Caution: summaries are a navigation aid, not a substitute. For anything contractual, regulatory, or finding-level, read the cited section yourself before you sign off.

07

NO-CODE

Internal Security Tools with Artifacts

From idea to working internal tool — in one conversation, zero dev tickets.

Claude Artifacts builds functional, styled web apps directly in the chat — live preview, iterate in plain English, export the HTML. For a security team, that means the small internal tools you've always wanted but never got engineering time for.

When to use it

- Phishing triage checklist app for the help desk — guided questions, verdict, escalation path
- Risk scoring calculators (likelihood × impact with your risk matrix baked in)
- Tabletop exercise injects — a clickable scenario walker for incident drills
- Security awareness quizzes and onboarding one-pagers
- Prototypes to show engineering exactly what you want before writing a ticket

COPY-PASTE PROMPT · PHISHING TRIAGE TOOL

```
You are an expert internal-tools designer who builds clean, fast web apps
for security teams.

Build a single-page phishing email triage tool for our help desk as an
artifact.

Flow: analyst answers 8 guided questions (sender domain age, link/domain
mismatch, urgency language, attachment type, requests credentials, etc.),
the tool computes a risk verdict (Benign / Suspicious / Malicious), and
shows the matching playbook step and escalation contact.

Navy blue and white theme, large touch-friendly controls, a "copy summary
to ticket" button. No external network calls, no data stored.
```

SECURITY NOTE FOR ARTIFACT TOOLS

- Keep these tools client-side and data-free by default — no secrets, no real user data baked in.
- Anything that will handle production data or authentication goes through normal SDLC and review, not an artifact.

08

CONNECTORS

Connectors for the SecOps Workflow

Plug Claude into the apps where your work already lives — with guardrails.

Connectors let Claude read and act on live data in tools like Gmail, Google Calendar, Slack, and Notion (enable under Settings → Connectors). Instead of copy-pasting context, Claude pulls it directly — your inbox, your meeting schedule, your team channel.

When to use it

- Inbox triage — flag the advisories, auditor requests, and vendor disclosures that actually need you today
- Meeting prep — pull the latest thread, related calendar invites, and open items in one shot
- Calendar defense — find conflicts with the on-call rotation, suggest reschedules
- Channel summaries — “what did #security-alerts decide while I was out?”

COPY-PASTE PROMPT · MORNING TRIAGE

```
You are my security operations chief of staff.

Using the Gmail connector, review my unread mail from the last 24 hours.

Surface, in priority order: (1) anything from auditors, regulators, or
legal; (2) vendor security advisories affecting [our stack]; (3) internal
escalations or incident threads; (4) everything else in one line each.

For the top 3, draft a short reply for my review. Do not send anything —
show me the drafts first.
```

CONNECTOR GROUND RULES FOR A SECURITY TEAM

- Least privilege applies to AI too: connect only the accounts and scopes you need.
- Claude asks before sending mail or taking consequential actions — keep it that way. Review drafts; you own what gets sent.
- Email content is untrusted input. If a message contains text that looks like instructions to the AI, Claude should — and will — surface it rather than obey it. Report anything odd.
- Check your organization’s AI policy before connecting mailboxes that contain regulated or customer data.

09

AUTOMATION

Co-Work & Claude Code

Beyond chat: Claude works on your machine — organizing, scripting, assembling.

Everything through Module 08 was chat: you type, Claude replies. Co-Work (in the Claude desktop app) and Claude Code (terminal) are agents: they take scoped access to folders and actually do the work — rename, sort, convert, script, assemble. For a security team, that's hours of evidence wrangling and report assembly off your plate.

Security-flavored use cases

- Organize an evidence folder after an incident — consistent naming, timeline order, manifest with file hashes
- Assemble the monthly report — pull the latest exports, normalize formats, produce the .docx
- Batch work — convert 200 screenshots for an audit binder, dedupe a shared drive folder
- Write and run small scripts — "parse these firewall configs and list any ANY-ANY rules"

COPY-PASTE PROMPT · EVIDENCE FOLDER ORGANISATION

```
You are a meticulous digital forensics assistant who organises case
material into clean, auditable structures.

Take access to the folder ~/Cases/INC-2026-041, scan everything inside,
and reorganise it into: /Timeline, /Logs, /Screenshots, /Comms,
/Reports, /Misc.

Rename files to YYYY-MM-DD_description format where dates are knowable.
Do NOT modify file contents. Generate a manifest.csv listing every file,
its new path, size, and SHA-256 hash.

Ask before deleting or overwriting anything. Summarise every move at the
end.
```

SAFETY NOTES — NON-NEGOTIABLE FOR CASE MATERIAL

- Never point an agent at original evidence — work on verified copies; chain of custody comes first.
- Start with low-stakes folders (Downloads, screenshots) to build trust in the workflow.
- Always require confirmation before deletes. Keep backups of anything irreplaceable.
- Agent runs are transparent — read what it did. You sign off on the result, not the tool.

10

THE SHORTCUT LAYER

/god · /ghost · /artifacts · /blindspot · /deepdive

Five reusable prompt shortcuts to standardize across the team.

Straight talk first: apart from Artifacts (a real Claude feature), these are **not** official built-in commands — you may see them hyped online as "secret modes." They work because Claude reads them as shorthand for an instruction. That's actually better for you: you get to define exactly what each one means, agree on it across your team, and get consistent behavior. Paste the definitions below into your user preferences or a Project's instructions (or define them as real custom slash commands in Claude Code), and the shortcuts become reliable.

`/god`	**Expert mode — peer-level depth, zero hand-holding**

"When I type /god: respond as a principal security engineer talking to a peer. Assume fluency with networking, identity, cloud, and detection engineering. Skip basics and definitions, give the strongest technical answer, state trade-offs and failure modes, and say 'unknown' rather than hedge." Use for: architecture reviews, detection logic, hard troubleshooting.

`/ghost`	**Ghostwriter — sounds like me, ready to send**

"When I type /ghost: write in my voice — plain, direct, no AI-isms, no 'I hope this finds you well.' Match the audience I name (exec, auditor, all-staff, engineer). Output only the final text, ready to paste, no commentary." Use for: incident comms, audit responses, policy announcements, awkward vendor emails.

`/artifacts`	**Build it — real, usable output**

"When I type /artifacts: don't describe it, build it. Produce the deliverable as an artifact or file — interactive tool, dashboard, document, or deck — and keep iterating on the same artifact as I give feedback." Use for: triage tools (Module 07), dashboards (Module 04), decks and docs (Module 05). This one maps to a genuine product feature.

`/blindspot`	**Red-team my thinking**

"When I type /blindspot: stop agreeing with me. Attack the plan or analysis above: what am I missing, what assumption is weakest, how would an attacker / an auditor / Murphy's law break this? Rank the top 5 blind spots by likelihood × impact and suggest one mitigation each." Use before: change windows, incident closure, control design sign-off, big purchases. The most security-native shortcut of the five.

`/deepdive`	**Exhaustive layer-by-layer analysis**

"When I type /deepdive: go exhaustive. Work the topic layer by layer — background, mechanics, root cause, second-order effects, open questions — and cite sources where facts are involved (use web search or Research mode if needed). Length is no constraint; completeness is." Use for: root-cause analysis, unfamiliar CVE classes, framework gap assessments.

11

GOVERNANCE

Safe Use of AI on a Security Team

Security teams have to model the behavior they ask of everyone else.

Data handling

- Know the data classes approved for AI tools in your organization — and the ones that never go in: credentials, keys, customer PII, regulated data, anything under legal hold.
- Sanitize before upload: hash or strip identifiers from logs, tickets, and evidence excerpts.
- Prefer org-managed accounts (Team/Enterprise) over personal accounts for work content.
- Connectors and agents get least privilege: minimum scopes, scoped folders, work on copies.

Verification

- Treat AI output as a competent first draft from a junior analyst: review like you'd review their work.
- Independently verify anything actionable — IOCs, CVE details, config and firewall changes, legal or compliance interpretations.
- Ask for citations and confidence levels; Research mode and web search provide sources you can check.

Prompt injection awareness

When Claude reads external content — web pages, emails, documents, scraped chatter — that content can contain text designed to manipulate AI tools. Claude treats observed content as data, not instructions, and will surface suspicious embedded instructions rather than follow them. Your part: notice when something odd appears in a workflow, and report it like you'd report any social-engineering attempt — because that's what it is.

A TEAM OPERATING AGREEMENT WORTH ADOPTING

- One shared Project with your team's prompt library, the shortcut definitions from Module 10, and a sanitization checklist.
- Every workflow gets a named owner who has run it at least three times before the team relies on it.
- Anything customer-facing or change-producing gets a human reviewer. No exceptions.
- Monthly 30-minute share-out: one workflow win, one failure, one prompt improvement.

APPENDIX A · TOOLS AND LINKS DIRECTORY

Every tool. One page.

Core AI platforms

TOOL	WHAT FOR	WHERE
Claude	Primary AI — chat, Research mode, Artifacts, Skills, Connectors	claude.ai
Claude Desktop	Required for Co-Work and local file access	claude.ai/download
Claude Code	Terminal agent for scripting and automation; supports custom slash commands	claude.ai/code
Grok	Live X search — useful raw OSINT signal	grok.com
Perplexity	Research-first AI with live citations	perplexity.ai

Claude features used in this guide

FEATURE	WHAT IT UNLOCKS	HOW TO ENABLE
Research Mode	Deep, cited, multi-source research (Modules 02, 03)	Toggle in claude.ai
Skills (pptx/docx/xlsx/pdf)	Real file deliverables (Module 05)	Automatic — just ask for the format
Artifacts	Interactive tools, dashboards, pages (Modules 04, 07)	Built into claude.ai
Connectors	Gmail, Calendar, Slack, Notion, and more (Module 08)	Settings → Connectors
Projects	Shared instructions + prompt library for the team	claude.ai → Projects

Verification sources — always check the primary record

TOOL	WHAT FOR	WHERE
NVD / MITRE CVE	Authoritative CVE details	nvd.nist.gov · cve.org
CISA KEV	Known exploited vulnerabilities catalog	cisa.gov/kev
MITRE ATT&CK;	Technique mapping for intel and detections	attack.mitre.org
Vendor PSIRTs	Ground-truth advisories for your stack	vendor security pages

“You’ve seen how to become a faster version of this team by orchestrating these tools well — and verifying everything they hand back. Pick one module. Run it Monday.”