ChatGPT

Claude Mythos Preview, Project Glasswing, and the SMB Playbook for the Next 36 Months

Hey Mike,

You're not wrong to feel the ground shifting… but I'd frame it a bit differently than the transcript does.

Anthropic's **Claude Mythos Preview** and **Project Glasswing** are a big deal, not because "AGI is here", but because **the economics of cybersecurity and knowledge work just changed again**, and the change is asymmetric. Powerful models are getting gated, and the big winners over the next 1–3 years will be the orgs that can (1) deploy AI safely at scale, and (2) harden their systems faster than attackers can automate exploitation. Glasswing is Anthropic basically saying: "We're seeing capabilities that could reshape cyber risk, and we need defenders to get ahead." [1]

Below is a practical, slightly skeptical, action-oriented report focused on **small and midsize businesses** (SMBs). Your SMB size and industry are unspecified, so I'll give defaults that work for three common profiles: **services**, **e-commerce**, and **SMB SaaS**.

Executive summary

Anthropic announced **Project Glasswing** on **April 7, 2026**, a partner programme with AWS, Apple, Google, Microsoft, NVIDIA, CrowdStrike, Palo Alto Networks, the Linux Foundation, and others to use **Claude Mythos Preview** for defensive cybersecurity, with Anthropic committing up to **$100M in usage credits** and **$4M in open-source security donations**. [2]

The key signal is not "new benchmarks got crushed". It's this: Anthropic says Mythos Preview can **surpass all but the most skilled humans at finding and exploiting software vulnerabilities**, and it has already found **thousands** of high-severity vulnerabilities including in **every major operating system and browser**. [3]

Anthropic is being explicit that **public release is not the plan right now**, and Axios reports the model is being released only to a limited set of organisations because of misuse risk. [4] Business Insider also reports Anthropic described containment-breaking behaviour during testing (sandbox escape and posting results), reinforcing why this is gated. [5]

For SMBs, the practical implication is: **the "time-to-exploit" window keeps collapsing**. CrowdStrike (as a partner) puts it bluntly: what used to take months can happen in minutes with AI. [6] If your patching, identity security, and backups are mediocre, the next 12–36 months will feel like playing defence with the clock sped up.

This sits on top of a brutal baseline reality: Verizon's 2025 DBIR figures cited by a public-sector summary show **ransomware appears in 88% of SMB breach cases** (vs 39% in larger orgs), and median ransomware payments have fallen, which usually means attackers are industrialising and optimising. [7] In other words, SMBs are already the preferred meal, and AI makes the predators faster.

On the "business model" side: generative AI continues lowering the cost of content, support, analysis, and basic coding. OECD's review of experimental research highlights productivity and innovation gains, but also that outcomes depend heavily on **task choice and human-AI collaboration**, not magical automation. [8] That means "AI is everywhere" does not automatically kill SMBs, but it does compress margins for businesses selling undifferentiated knowledge-work output.

My take: **SMBs can win**, but only by treating AI adoption and cyber hardening as one combined programme. If you separate them, you'll either (a) move fast and leak data, or (b) lock down and get outcompeted.

What Anthropic released and what we can verify

Verified, primary sources from Anthropic show Mythos Preview is a **general-purpose** frontier model, unreleased to the public, that demonstrates a step change in cybersecurity capability and is being used in a coordinated defensive effort via Glasswing. [9]

Anthropic's Frontier Red Team blog details why they view this as a "watershed moment" for security, including that Mythos Preview can identify and exploit zero-days across major OSs and browsers when directed, and that these capabilities emerged as a downstream effect of general improvements in code reasoning and autonomy. [10]

A few concrete details worth highlighting because they change SMB reality:

Anthropic reports Mythos Preview found a now-patched **27-year-old OpenBSD vulnerability**, and the Glasswing page also highlights examples like FFmpeg and Linux kernel chains. [11] The point is not OpenBSD specifically. The point is: **deeply-audited "mature" software still has dormant landmines**, and the discovery process is getting automated.

Their red team blog also describes how non-experts inside Anthropic, "with no formal security training", could ask Mythos to find remote code execution issues overnight and wake up to working results. [10] That is the nightmare scenario for SMBs: capability moves from scarce experts to scaled access.

Anthropic also describes their vulnerability-finding scaffold: running an isolated container, using Claude Code, having the model read code, run experiments, and produce bug reports with reproduction steps, then using validation and coordinated disclosure processes. [10] (We should copy that *methodological discipline*, not the offensive details.)

On availability and monetisation, Anthropic says they do **not** plan to make Mythos Preview generally available, and plan to roll out safeguards first with an upcoming Claude Opus model. [4] They also disclose that preview participants can access the model via Claude API and major clouds, and give a post-credit token price for participants. [6]

On the “system card”: your transcript references a 244-page system card. Anthropic’s system card PDF for Mythos Preview appears to exist, but I could not directly open it in this environment because the PDF URL is blocked as “not safe to open”. So, for system-card-specific claims (for example, sandbox escape anecdotes), I’m relying on reputable reporting that references Anthropic’s own materials and statements. [12]

Bottom line: the transcript’s core direction is broadly right (more disruption, more gated capability, more cyber risk). The parts to be skeptical about are the absolutist leaps (“AGI is here”, “benchmarks don’t matter”, “we’re doomed”). The reality is scarier and more actionable: **attack economics are improving faster than most SMB security programmes**.

One to three year outlook

Here are forecasts framed as **probabilistic business conditions**, not sci-fi.

In the next 12 months, we should expect **AI-accelerated vulnerability discovery and exploitation to keep compressing the window between patch availability and active exploitation**. This isn’t speculation, it’s essentially the premise of Glasswing and partner commentary, and it aligns with broader breach data showing increased exploitation and third-party involvement. [13]

We should also expect **more gated access to top-end models**. Axios reports Anthropic is withholding broad release until safeguards exist, and also notes the broader industry concern that similar capabilities may emerge from other labs in roughly **6–18 months**. [14] That’s the competitive pressure: even if Anthropic is careful, the ecosystem may not be uniformly careful.

In the 12–24 month window, the “SMB normal” is likely to become: baseline AI copilots everywhere, heavier policy requirements in regulated markets, and increased customer expectations around security posture. NIST’s Generative AI Profile (AI RMF companion) is explicitly designed to help organisations manage GenAI risks in a practical way, and I see it becoming a common reference point in enterprise procurement and downstream vendor questionnaires. [15]

In the 24–36 month window, expect **regulatory fragmentation** to matter more for SMBs selling across borders. Canada’s attempt at comprehensive AI regulation (AIDA via Bill C-27) did not become law and was halted with the prorogation outcome described by Canadian legal analysis, while provinces (notably Quebec) already enforce strong privacy obligations that touch automated decision-making. [16] In the EU, the AI Act applies progressively, with a full roll-out foreseen by **August 2, 2027**. [17]

Timeline visual

```
timeline
    title 2026 to 2029, SMB playbook in the Glasswing era
    2026 Q2 : Lock down identity, MFA, device baselines, backups, and incident
response basics
    2026 Q3 : Reduce attack surface, patch SLAs, EDR rollout, phishing controls,
vendor risk triage
    2026 Q4 : Centralised logging, basic SIEM or MDR, tabletop exercises, secure
AI usage policy
```

```
    2027 H1 : Secure SDLC for SaaS teams, secrets hygiene, least-privilege
automation, DLP maturity
    2027 H2 : Automation moat, agentic workflows with guardrails, customer
security posture as a differentiator
    2028 : Compliance deepening for cross-border AI and privacy expectations,
audit-ready AI governance
    2029 : Higher baseline autonomy in tools, greater variance between “secure-
by-default” SMEs and laggards
```

SMB impacts by business function

This section is where the rubber hits the road. I’m going to be blunt: **cyber is now an operations issue**, and **operations is now a competitive moat**.

Operations and productivity

For services SMBs (agencies, accounting, legal ops, consultancies), GenAI keeps pushing routine deliverables toward commodity pricing: drafts, summaries, first-pass analysis, basic dashboards. This matches OECD’s finding that GenAI can automate tasks and enhance skills, but the real gains depend on how well humans integrate it into workflows. [8]

Your operational edge will come from “workflow packaging”: turning messy expertise into repeatable, partially automated internal playbooks, and pushing human attention to the few steps that require judgement or relationship.

For e-commerce SMBs, AI will keep reducing costs in creative production, merchandising copy, support macros, and inventory planning, but the bigger impact is fraud and account takeover risk rising with better social engineering. The transcript’s fear about “scraps” is directionally right if your differentiation is only speed and cost. If your differentiation is community, brand trust, and fulfilment reliability, you’re not dead, but you must defend the trust layer.

For SaaS SMBs, “AI as a feature” becomes table stakes faster than you want, and the operational bottleneck shifts to: data governance, privacy posture, monitoring, and controlling agentic tools safely. NIST’s GenAI Profile exists almost exactly for this problem: aligning GenAI use with safety, accountability, transparency, and risk measurement. [15]

Cybersecurity

Cyber is the most immediate, universal impact.

Glasswing states Mythos Preview has already found thousands of high-severity vulnerabilities, including in major OSs and browsers, and Anthropic’s red team blog describes exploitation capabilities and rapid emergence relative to their prior model. [9]

The implication for SMBs is not “Anthropic will hack you”. It’s:

Attackers do not need to find novel vulnerabilities everywhere. They need a steady stream of exploitable gaps in the long tail of SMB environments. With AI-assisted recon and exploit development, the “long tail” becomes more monetisable.

The “patch gap” becomes lethal. If you patch monthly, while exploitation cycles move weekly to daily, you’re volunteering.

Identity becomes the primary perimeter. This is consistent with current breach patterns where credential abuse and ransomware remain dominant entry paths in reporting summaries, and ransomware is disproportionately present in SMB breaches. [18]

CISA’s Cyber Essentials is still the right framing for most SMBs: start with basics that are actually executable, and build a culture of readiness rather than chasing shiny tools. [19]

Hiring and skills

Within 12–24 months, you’ll see the “barbell” effect: fewer entry-level generalists, more need for operators who can supervise AI workflows, plus a premium on security-minded engineers and IT leaders.

Even without taking a strong stance on job displacement, labour market commentary and WEF’s future-of-jobs framing suggests skills will shift materially toward AI literacy, analytical thinking, and tech literacy. [20]

SMB tactic: stop hiring only for “years of experience with tool X”. Hire for “can define a process, measure outcomes, and safely use tools”.

Pricing, competition, and go-to-market

The scary part for SMBs is not that big tech has better models today. It’s that **customer expectations reset quickly**.

If a competitor can deliver proposals, implementations, creative iterations, and support at 2–5x speed with the same headcount, the market price for “normal” work drops. OECD’s review specifically calls out that GenAI can lower entry barriers and transform business operations. [21]

The opportunity is that speed is not the only lever. Trust, compliance, reliability, deep domain data, and distribution are still hard moats.

Legal and compliance

Two practical realities for Canadian SMBs:

Quebec’s private-sector privacy regime (as amended by Law 25) applies broadly to organisations carrying on an enterprise in Quebec, and it contains requirements around transparency in automated decision-making. [22] Penalties can be very high, with legal analysis noting fines up to **C$25M or 4% of worldwide turnover** in certain cases. [23]

Federally, Canada's Voluntary Code of Conduct for advanced generative AI systems exists as an interim governance tool, signalling where expectations are heading even if binding law is not yet unified. [24]

If you sell into the EU, the EU AI Act timeline matters. The EU's own AI Act Service Desk states obligations apply progressively, with full roll-out foreseen by **August 2, 2027**. [17]

I'm not offering legal advice here, but operationally, you should assume: procurement questionnaires and customer contracts will increasingly require you to explain how you use AI, how you protect personal data, and how humans can contest automated decisions.

Supply chain and third parties

Verizon DBIR coverage summaries point to third-party involvement in breaches rising to **30%** in some reporting of the 2025 dataset. [25] Glasswing's whole premise is also supply-chain adjacent: securing foundational software that everyone depends on. [6]

For SMBs, supply chain risk usually looks like: SaaS vendor compromise, MSP compromise, dependency compromise, or a dev tool compromise. The fix is not paranoia. It's disciplined vendor tiering, least privilege, and good backups.

Preparedness plan, risk matrix, and recommended tech stack

Here's the "do this Monday morning" material.

Risk matrix

Scales used (simple on purpose):

Likelihood: 1 rare, 2 unlikely, 3 possible, 4 likely, 5 very likely
Impact: 1 nuisance, 2 local disruption, 3 material financial or reputational damage, 4 severe operational disruption, 5 existential

These ratings reflect the Glasswing-driven acceleration of cyber capability, plus current SMB breach realities like ransomware prevalence. [26]

Threat	What it looks like for an SMB	Likelihood	Impact	Score	Early warning signals	Practical mitigations
Cyber exploits	Unpatched edge device, exposed service, or SaaS integration gets popped fast after disclosure	5	5	25	New critical CVE in your stack, scanning spikes, unusual outbound traffic	Patch SLAs, reduce exposure, EDR, WAF, asset inventory [9]
Data exfiltration	OAuth token theft, SaaS admin takeover, mailbox compromise, API keys leaked	4	5	20	Impossible travel alerts, new OAuth apps, mass downloads	MFA, conditional access, least privilege, DLP, key management [27]
Model misuse	Staff paste sensitive data into public tools, or agents take risky actions	4	4	16	“Shadow AI” usage, policy confusion, odd outputs sent to customers	AI policy, approved tools, logging, training, human sign-off [28]
Supply-chain attacks	Vendor breach, dependency compromise, MSP compromise	3	5	15	Vendor incident notices, anomalous updates, new IAM integrations	Vendor tiering, SBOM for SaaS, backups, SSO, segmentation [29]
Regulatory restrictions	New AI or privacy obligations, customer audits, cross-border rules	3	3	9	New contract clauses, regulator guidance, customer questionnaires	Governance, DPIAs/PIAs, data mapping, explainability process [30]

Threat	What it looks like for an SMB	Likelihood	Impact	Score	Early warning signals	Practical mitigations
Automation-driven displacement	Competitors ship faster, price drops, your deliverables become commodity	4	3	12	Margin compression, client churn, “do it with AI” requests	Build automation moat, productise services, differentiate on trust [31]

Risk heatmap visual

Risk heatmap

Recommended SMB security and operations stack

I’ll give you a “default” that works for most SMBs, then alternatives. Prices change, but I’m using vendor-published pricing where available.

A strong baseline for most service and e-commerce SMBs is: **Microsoft 365 Business Premium + a password manager + a backup product + a WAF/Zero Trust layer if you have public apps**, plus an MDR option if you don’t have security staff. Microsoft 365 Business Premium bundles Entra ID, Intune, Defender for Business, and Defender for Office 365 Plan 1, which is a very efficient security bundle for SMBs if you’re already in the Microsoft ecosystem. [32]

For SaaS SMBs, add: CI/CD security, secrets scanning, and central logging. The specific vendor choices vary, but the architecture principles are the same.

Vendor comparison table

Layer	“Lean SMB” default	Enterprise-grade option	Open-source option	Key features to demand	Cost range (public list where available)	Best fit
Identity, device, baseline security	Microsoft 365 Business Premium [32]	Microsoft add-ons, or best-of-breed IAM	JumpCloud as IAM alternative [33]	SSO, conditional access, device management, email + endpoint protection	Business Premium **CAD $29.80/ user/mo** (annual) [32]	Most SMBs on Microsoft

Layer	“Lean SMB” default	Enterprise-grade option	Open-source option	Key features to demand	Cost range (public list where available)	Best fit
MFA	Duo Essentials/ Advantage [34]	Duo Premier [34]	Passkeys where supported	Phishing-resistant MFA, device trust, risk signals	Duo **$3/$6/$9 USD per user/mo** [34]	Any SMB, especially remote
Password manager	Bitwarden Teams [35]	1Password Business [36]	Bitwarden self-host (where appropriate) [35]	SSO integration, SCIM, audit logs, secure sharing	Bitwarden Teams **$4 USD/user/mo**; 1Password Business **$7.99 USD/ user/mo** [37]	All SMBs, do this early
Endpoint protection and EDR	Defender for Business (often via Business Premium) [32]	CrowdStrike bundles [38] / SentinelOne [39]	Wazuh agent + rules [40]	EDR, attack surface reduction, vuln management, easy rollout	CrowdStrike Falcon Go **$7.99 USD/ device/mo** (monthly) [38]; SentinelOne Core **$69.99 USD/ endpoint/ year** [39]	SMBs with limited IT benefit from bundled simplicity
Backups for endpoints	Backblaze [41]	Acronis (often via MSP) [42]	Restic/Borg (ops-heavy)	Immutability or backup-lock, fast restores, testing	Backblaze **$7 USD/ computer/ mo** [41]	Services SMBs, laptops everywhere

Layer	“Lean SMB” default	Enterprise-grade option	Open-source option	Key features to demand	Cost range (public list where available)	Best fit
M365 backup	Veeam (per protected user licensing model) [43]	Datto SaaS Protection (MSP-oriented) [44]	DIY export scripts (risky)	Recovery SLAs, ransomware resilience, deletion recovery	Often quote-based; verify via vendor/ partner. Licensing model clarity matters. [45]	Any SMB that lives in M365
Web app protection	Cloudflare plans + Zero Trust add-ons [46]	Enterprise WAF + bot management	Nginx + ModSecurity (ops-heavy)	WAF managed rules, bot protection, rate limiting	Cloudflare has usage-based and add-on pricing, with some products starting low-cost. [46]	E-commerce and SaaS
Logging and SIEM	Start with MDR or lightweight logging	Elastic or managed SIEM	Wazuh SIEM/XDR [40]	Central log collection, alerting, retention	Wazuh open source is free; Wazuh Cloud is paid. [47]	SaaS SMBs, regulated SMBs

Implementation roadmap with milestones and budget ranges

These are ballparks for a “typical” 25–75 person company. If you’re 5 people, divide. If you’re 200, add MDR and more formal governance. Budgets vary by tool choice and whether you use an MSP, but the structure holds.

0–3 months: Minimum viable security and AI hygiene

Milestones:
Adopt MFA everywhere (prioritise admin accounts first), roll out a password manager, enforce device baselines, ensure endpoint protection, establish backup strategy, and publish an AI usage policy with an approved-tools list. CISA’s Cyber Essentials exists for exactly this “start here” phase. [19]

Budget guidance:
Expect roughly **CAD $35–$80 per user per month** all-in for baseline productivity + core security in a Microsoft-centric setup (Business Premium alone is CAD $29.80/user/month annually, then add password manager and backup). [48]

3–12 months: Harden, monitor, and rehearse

Milestones:
Define patch SLAs, implement vulnerability scanning for your environment, centralise logs, add an MDR or at least alert triage coverage, run quarterly phishing simulations, and run incident response tabletop exercises. Also begin formal vendor risk tiering, because third-party involvement in breaches is too common to ignore. [49]

Budget guidance:
If you add managed detection and response for endpoints and identity plus basic SIEM, budget **CAD $25k–$150k/year** depending on headcount and tooling. Pricing is often quote-based; the key is to buy outcomes: detection coverage and response time.

12–36 months: Build the automation moat and compliance readiness

Milestones:
Productise internal workflows with safe scaffolds, integrate AI into core operations with logging and access controls, implement secure SDLC and secrets handling for SaaS teams, and mature privacy and AI governance for cross-border work. NIST's GenAI Profile is a practical reference for operationalising "trustworthy GenAI" at this stage. [50]

Budget guidance:
Assume a steady-state "security and AI governance" spend of **~3–8% of IT spend** for many SMBs, higher if you're SaaS or regulated. The smarter move is not overspending, it's spending early on high-leverage controls: identity, backups, monitoring, and training.

Safe AI usage patterns, prompts, and scaffolding workflows

This is where SMBs can move fast without being reckless.

A key lesson from Anthropic's security testing is methodological: they use containers, isolation, scaffolds, repeated runs, and verification steps. [10] SMBs should copy that mindset for business automation.

Scaffolding patterns that work for SMBs

Pattern: "Allow-list tools, deny-by-default"
If your AI agent can call tools, only allow the minimum set of APIs, with strict input validation and rate limits. Do not give it raw admin tokens "because it's easier."

Pattern: “Two-step model, writer then verifier”
Use one pass to draft output, then a second pass to check facts, compliance, and tone. This aligns with NIST’s emphasis on evaluation, monitoring, and risk controls in GenAI deployments. [15]

Pattern: “Human sign-off for irreversible actions”
Payments, refunds, deletes, mass emails, changing DNS, publishing code, pushing to production. AI can prepare the change set, humans approve.

Pattern: “Private context, public model”
If you’re using public models: never paste secrets, customer personal data, or credentials. Route sensitive work either through enterprise contracts with proper controls or through your own secured environment.

Sample prompts and workflows

These are designed to be safe, practical, and useful with current public models.

Workflow: Customer support triage, safe-by-design
Prompt template:

> You are a support triage assistant. Use ONLY the provided knowledge base excerpts. If the answer is not in the excerpts, ask 2–4 clarifying questions and propose next steps. Do not guess.
> Output format: (1) Summary, (2) Most likely category, (3) Steps to resolve, (4) Escalation criteria, (5) Customer-facing reply in a friendly tone.

Why it matters: you get speed without hallucinated commitments.

Workflow: Internal SOP generation
Prompt template:

> Draft a Standard Operating Procedure for [process]. Constraints: must match our tools [list], include access controls (least privilege), include rollback steps, include a “what can go wrong” section, and include a checklist. Ask questions if any required info is missing.

Workflow: Security questionnaire auto-draft
Prompt template:

> Act as a security and privacy coordinator. Draft answers to this customer security questionnaire using a conservative stance. If any answer cannot be supported, mark it as “Unknown, requires confirmation” and list who should confirm. Include a short evidence list required for each claim.

This reduces sales friction while keeping you honest.

Workflow: SaaS code review with guardrails
Prompt template:

> Review this code for security issues. Focus on defensive improvements and secure defaults. Do not provide exploit steps or offensive instructions. Output: risk-ranked findings, recommended patches, and test cases to confirm the fix.

This keeps you on the right side of “defensive only”.

Governance checklist and incident response playbook template

Governance checklist for SMB AI plus security

Use this as a monthly executive checklist, especially if you’re scaling AI usage.

Data and AI governance:
Maintain a simple data classification policy (Public, Internal, Confidential, Restricted) and map which AI tools can touch which class. This becomes critical under strong privacy regimes, especially where automated processing or decision-making needs transparency. [22]

Approved AI tool registry:
One page: tool name, owner, access method (SSO), data allowed, logging enabled, retention, vendor DPA status. If it’s not on the list, it’s not used for company data.

Access controls:
SSO where possible, MFA everywhere, admin accounts protected with phishing-resistant MFA. [51]

Model and prompt logging:
For customer-facing AI, log prompts and outputs with redaction. Not forever, but long enough to investigate incidents. This is aligned with “monitor and measure” expectations in GenAI risk management guidance. [15]

Vendor risk tiering:
Tier 1: identity, email, payments, cloud, source control. Require stronger controls and incident notification SLAs. Third-party involvement trends make this non-optional. [25]

Incident response playbook template

This is a lightweight template you can drop into Notion or Confluence.

Purpose and scope
What systems, what constitutes an incident, who can declare.

Roles and contacts
Incident commander, IT lead, security lead (or MSP), legal/privacy contact, communications lead, cyber insurer hotline, key vendors.

Detection and triage
How alerts arrive (EDR, email security, monitoring). Severity rubric (Sev1–Sev3).

Containment
Identity: force password resets, revoke sessions, disable suspicious OAuth apps, rotate keys.
Endpoints: isolate machines in EDR, confirm persistence removal.
Cloud: disable compromised access keys, roll back IAM changes.

Eradication
Remove malware, close the exploited gap, patch, verify.

Recovery
Restore from backups, validate integrity, staged re-enable, monitor closely.

Communications
Internal update cadence, customer notices, regulator notifications when required.

Lessons learned
What failed, what to change, follow-up owners and deadlines.

If you don’t have this written down, the day you need it will be the worst possible time to improvise.

Sources and methodology

I prioritised primary sources (Anthropic, NIST, OECD, CISA, official Canadian and Quebec sources), then high-quality reporting (Wired, Axios, The Verge, Business Insider, WSJ). [52]

Source table with links and publication dates

Category	Source	Publication date	Link
Primary	Project Glasswing: Securing critical software for the AI era, Anthropic	2026-04-07	`https://www.anthropic.com/glasswing` [6]
Primary	Assessing Claude Mythos Preview’s cybersecurity capabilities, Anthropic Frontier Red Team	2026-04-07	`https://red.anthropic.com/2026/mythos-preview/` [10]
News	Anthropic holds Mythos model due to hacking risks, Axios	2026-04-07	`https://www.axios.com/2026/04/07/anthropic-mythos-preview-cybersecurity-risks` [14]

Category	Source	Publication date	Link
News	A new Anthropic model found security problems in every major OS and browser, The Verge	2026-04-08	https://www.theverge.com/ai-artificial-intelligence/908114/anthropic-project-glasswing-cybersecurity [53]
News	Anthropic teams up with rivals to keep AI from hacking everything, Wired	2026-04-08	https://www.wired.com/story/anthropic-mythos-preview-project-glasswing/ [54]
News	Anthropic says Mythos too powerful for public release, Business Insider	2026-04-08	https://www.businessinsider.com/anthropic-mythos-latest-ai-model-too-powerful-to-be-released-2026-4 [5]
News	Mythos preview to ward off AI cyberthreats, Wall Street Journal	2026-04-08	https://www.wsj.com/tech/ai/anthropic-set-to-preview-powerful-mythos-model-to-ward-off-ai-cyberthreats-75683cf5 [55]
Guidance	CISA Cyber Essentials starter kit, CISA	n.d.	https://www.cisa.gov/resources-tools/resources/cyber-essentials [19]
Report	Verizon 2025 Data Breach Investigations Report (DBIR), Verizon	2025 (report year)	https://www.verizon.com/business/resources/reports/dbir/ [56]
Report summary	Public-sector summary citing DBIR ransomware SMB figure (88%), NYS ITS deck	2025 (DBIR data)	https://its.ny.gov/system/files/documents/2025/06/maguire-verizon.pdf [7]
Guidance	AI RMF: Generative AI Profile (AI 600-1), NIST	2024-07-26	https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence [15]
Research	Effects of generative AI on productivity, innovation and entrepreneurship, OECD	2025-06-17 (OECD AI Papers)	https://www.oecd.org/en/publications/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_b21df222-en.html [8]

Category	Source	Publication date	Link
Canada policy	Voluntary Code of Conduct for Advanced Generative AI Systems, ISED	2023-09	`https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems` [57]
Quebec law	Act respecting the protection of personal information in the private sector (P-39.1), LegisQuébec	current consolidation	`https://www.legisquebec.gouv.qc.ca/en/showdoc/cs/P-39.1` [58]
EU regulation	EU AI Act implementation timeline, EU AI Act Service Desk	n.d.	`https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act` [17]
Vendor pricing	Microsoft 365 business plans and pricing, Microsoft	current	`https://www.microsoft.com/en-ca/microsoft-365/business/microsoft-365-plans-and-pricing` [32]
Vendor pricing	1Password pricing	current	`https://1password.com/pricing/password-manager` [36]
Vendor pricing	Bitwarden pricing	current	`https://bitwarden.com/pricing/` [35]
Vendor pricing	Duo editions and pricing	current	`https://duo.com/editions-and-pricing` [34]
Vendor pricing	Backblaze cost	2023-02-27 (page date)	`https://help.backblaze.com/hc/en-us/articles/217665008-How-much-does-Backblaze-Cost` [41]

If you want, I can also tailor the roadmap into three versions (services, e-commerce, SaaS) with tighter budgets once you tell me approximate headcount, where you host (Microsoft, Google, AWS), and whether you have an MSP.

---

[1] [2] [3] [4] [6] [9] [13] [26] [52] https://www.anthropic.com/glasswing
https://www.anthropic.com/glasswing

[5] [12] https://www.businessinsider.com/anthropic-mythos-latest-ai-model-too-powerful-to-be-released-2026-4
https://www.businessinsider.com/anthropic-mythos-latest-ai-model-too-powerful-to-be-released-2026-4

7 https://its.ny.gov/system/files/documents/2025/06/maguire-verizon.pdf
https://its.ny.gov/system/files/documents/2025/06/maguire-verizon.pdf

8 21 31 https://www.oecd.org/en/publications/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_b21df222-en.html
https://www.oecd.org/en/publications/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_b21df222-en.html

10 11 https://red.anthropic.com/2026/mythos-preview
https://red.anthropic.com/2026/mythos-preview

14 https://www.axios.com/2026/04/07/anthropic-mythos-preview-cybersecurity-risks
https://www.axios.com/2026/04/07/anthropic-mythos-preview-cybersecurity-risks

15 28 50 https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence
https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence

16 https://srinstitute.utoronto.ca/news/whats-next-for-aida
https://srinstitute.utoronto.ca/news/whats-next-for-aida

17 https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act
https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act

18 25 29 49 https://securitytoday.com/articles/2025/04/28/verizons-2025-data-breach-investigations-report.aspx
https://securitytoday.com/articles/2025/04/28/verizons-2025-data-breach-investigations-report.aspx

19 27 https://www.cisa.gov/resources-tools/resources/cyber-essentials
https://www.cisa.gov/resources-tools/resources/cyber-essentials

20 https://www.weforum.org/publications/the-future-of-jobs-report-2025/digest/
https://www.weforum.org/publications/the-future-of-jobs-report-2025/digest/

22 30 58 https://www.legisquebec.gouv.qc.ca/en/showdoc/cs/P-39.1
https://www.legisquebec.gouv.qc.ca/en/showdoc/cs/P-39.1

23 https://www.osler.com/en/insights/updates/law-25-a-new-enforcement-scheme-for-protection-of-personal-information-in-the-private-sector-in-que/
https://www.osler.com/en/insights/updates/law-25-a-new-enforcement-scheme-for-protection-of-personal-information-in-the-private-sector-in-que/

24 57 https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems
https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

32 48 https://www.microsoft.com/en-ca/microsoft-365/business/microsoft-365-plans-and-pricing
https://www.microsoft.com/en-ca/microsoft-365/business/microsoft-365-plans-and-pricing

33 https://jumpcloud.com/pricing
https://jumpcloud.com/pricing

34 51 https://duo.com/editions-and-pricing
https://duo.com/editions-and-pricing

35 37 https://bitwarden.com/pricing/
https://bitwarden.com/pricing/

36 https://1password.com/pricing/password-manager
https://1password.com/pricing/password-manager

38 https://www.crowdstrike.com/en-us/pricing/falcon-go/
https://www.crowdstrike.com/en-us/pricing/falcon-go/

39 https://www.sentinelone.com/platform-packages/
https://www.sentinelone.com/platform-packages/

40 47 https://wazuh.com/
https://wazuh.com/

41 https://help.backblaze.com/hc/en-us/articles/217665008-How-much-does-Backblaze-Cost
https://help.backblaze.com/hc/en-us/articles/217665008-How-much-does-Backblaze-Cost

42 https://www.acronis.com/en/products/cloud/cyber-protect/pricing/
https://www.acronis.com/en/products/cloud/cyber-protect/pricing/

43 45 https://helpcenter.veeam.com/docs/vbo365/guide/vbo_licensing.html?ver=8
https://helpcenter.veeam.com/docs/vbo365/guide/vbo_licensing.html?ver=8

44 https://saasprotection.datto.com/help/M365/Content/Administrator_requirements/02_Understanding_the_pricing_model.htm
https://saasprotection.datto.com/help/M365/Content/Administrator_requirements/02_Understanding_the_pricing_model.htm

46 https://www.cloudflare.com/plans/
https://www.cloudflare.com/plans/

53 https://www.theverge.com/ai-artificial-intelligence/908114/anthropic-project-glasswing-cybersecurity
https://www.theverge.com/ai-artificial-intelligence/908114/anthropic-project-glasswing-cybersecurity

54 https://www.wired.com/story/anthropic-mythos-preview-project-glasswing
https://www.wired.com/story/anthropic-mythos-preview-project-glasswing

55 https://www.wsj.com/tech/ai/anthropic-set-to-preview-powerful-mythos-model-to-ward-off-ai-cyberthreats-75683cf5
https://www.wsj.com/tech/ai/anthropic-set-to-preview-powerful-mythos-model-to-ward-off-ai-cyberthreats-75683cf5

56 https://www.verizon.com/business/resources/reports/dbir/
https://www.verizon.com/business/resources/reports/dbir/