Technology E&O research: what I did and key results

What I did

- Wrote a focused search prompt in `prompts/TEO_Search.xml` to systematically find Technology E&O offerings (carriers + MGAs), artifacts (specimens/wordings, product pages), and SERFF references.
- Ran parallel web research: ChatGPT Agent Mode and Perplexity Deep Search to capture candidates and public artifacts.
- Consolidated and deduplicated into `analysis/tech_eo_catalog_merged.csv` (19 unique providers).
- For each policy, downloaded specimen/wording PDFs and relevant brochures; created a dedicated folder under `final/##_Carrier – Product` and stored sources there.
- Extracted each policy into a normalized YAML using `prompts/policy_extraction.XML` (and `prompts/policy_extraction_followup.XML` to resolve unknowns). Saved per-policy `output.yaml` alongside sources.
- Scored all `output.yaml` files with `scripts/score_policies.py`, producing `policy_scores.csv` / `policy_scores.xlsx`.
- Reviewed and verified the top results in `policy_scores.csv` against public sources; updated affected `output.yaml` where needed.
- Selected four finalists (not strictly the top 4 by score): Corvus, Embroker, Beazley, and AXA XL. Included CFC (provided by Corgi) in the comparison deck for founder context.
    - Built founder-facing slides: 1-page side-by-side comparison (4 + CFC) and 5-page deep-dives per selected policy.

process (with references)

1. Search design
    - Authored `prompts/TEO_Search.xml` with scope, synonyms, and stepwise site/PDF/SERFF searches; standardized outputs (CSV + Markdown) and recency rule (≤5 years preferred, legacy flagged).
    - Defined comparison schema upfront in `prompts/policy_extraction.XML` to normalize outputs (fixed keys: trigger, duty, defense_costs, consent_to_settle, limits, sublimits, BI flags, exclusions, definitions, conditions) for easy diffing across policies.
2. Parallel collection (past chats)
    - ChatGPT Agent run produced `deliverables/searches/chat/tech_eo_catalog.csv`.

- Perplexity Deep Search produced `deliverables/searches/perplexity/tech_eo_catalog.csv` and `deliverables/searches/perplexity/sources.md` (raw URLs + access dates).
- Focused on carriers/MGAs with public wordings/specimens and verifiable marketing pages; excluded generic explainers without identifiable products.

3. Dedupe and merge
   - Consolidated to `analysis/tech_eo_catalog_merged.csv` using a composite key (Carrier/MGA + Product name + at least one public link).
   - Tie-break rules: prefer specimen/wording link over marketing; prefer more recent artifacts; keep admitted/E&S detail when available; retain legacy forms with a note.
4. Repository organization
   - Created per-policy folders under `final/##_Carrier – Product` (sequential prefix for ordering). Used `scripts/rename_folders_with_prefix.py` to normalize naming.
   - Ensured presence of `output.yaml` in each folder and placed all source PDFs/brochures there (e.g., specimen wordings, coverage guides).
5. Structured extraction and follow-up
   - Extracted each policy into the standardized YAML via `prompts/policy_extraction.XML`.
   - When fields were `null`/"Not found", ran `prompts/policy_extraction_followup.XML` to resolve only unknowns using `policy.metadata.source.link` (no changes to confirmed fields).
6. Scoring and verification
   - Scored all policies with `scripts/score_policies.py` → `policy_scores.csv`, `policy_scores.xlsx`.
   - Score inputs (see CSV columns): `n_core`, `n_coverage`, `n_defs_carves`, `n_limits`, `n_ops`, plus scenario checks (e.g., `Scenario_SaaS_API`, `Scenario_UGC_Defa`). Aggregated into `Score_Default`, `Score_B2B_SaaS`, `Score_Consumer` with a simple bonus/penalty field.
   - Read the top 10 by default score and verified against public sources, updating `output.yaml` where citations or details needed correction.
7. Selection and presentation
   - Selected 4 finalists by practical founder relevance (ranks 1, 2, 3, and 5) rather than strict top-4.
   - Built a 1-page comparison (4 + CFC) and 5-page deep dives per selected policy, emphasizing insuring agreements, key exclusions/carve-backs, hammer, defense costs, limits/retentions, BI/system-failure/dependent-provider nuances, and notable endorsements.

What each output.yaml captures and why it matters

- Purpose: Normalize every policy into comparable building blocks so a founder, GC, or broker can line up forms side-by-side and see what truly changes risk, negotiation leverage, and claims outcomes.

```
policy:
  metadata: { carrier, product, form_code, edition_date, jurisdiction, source }
  compare: { trigger, duty, defense_costs, consent_to_settle: { required, hammer, detai
  limits: { per_claim_limit, aggregate_limit, retention, sublimits: [ { name, amount, a
  coverage:
    tech_services_eo: { covered, summary, cite }
    media_liability: { covered, summary, cite }
    ip_infringement: { covered, summary, cite }
    privacy: { covered, summary, cite }
    network_security: { covered, summary, cite }
    business_interruption: { covered, dependent_providers, system_failure, cite }
    regulatory: { covered, summary, cite }
    pci: { covered, summary, cite }
    breach_response: { covered, summary, cite }
    cyber_extortion: { covered, summary, cite }
    social_engineering: { covered, summary, cite }
    contractual_liability_carveback: { exists, summary, cite }
  exclusions: [ { label, effect, carvebacks, cite } ]
  definitions: { professional_services, technology_services, wrongful_act }
  conditions: { notice_reporting, extended_reporting, territory }
  unknowns: [ { item, searched, next_best_sources } ]
```

- metadata: Anchors provenance and recency.
  - carrier/product/form_code/edition_date/jurisdiction: Identifies the exact form and state regime. Edition changes often alter coverage; this prevents apples-to-oranges.
  - source: Verifiable link + citation so every statement can be traced for underwriting and claims.
- compare: Core mechanics that shape defense and settlement.
  - trigger: Claims-made vs claims-made-and-reported drives reporting discipline; missed windows are a common denial vector for lean teams.
  - duty: Duty-to-defend vs indemnity (or hybrid) defines who controls counsel and strategy.
  - defense_costs: Inside vs outside limits materially changes remaining indemnity—critical when total limits are ≤$2M.

    - consent_to_settle/hammer: Soft/hard hammer percentages reveal settlement friction and residual defense sharing if you refuse a carrier-recommended deal.
- limits: How much actually pays and where it's capped.
    - per_claim_limit / aggregate_limit: Satisfy contractual requirements and board risk appetite.
    - retention: First dollars at risk; finance needs this to budget and negotiate.
    - sublimits: Surface practical caps for first-party modules (incident response, BI/DBL, cyber extortion, regulatory, PCI, reputational harm, claim prep). Many founders assume full limits; most forms sublimit these.
- coverage: Mapped to common startup exposures.
    - tech_services_eo: Negligence and certain contract carve-backs for SaaS/MSP/integration—essential for B2B SLAs.
    - media_liability & ip_infringement: Defamation and copyright/trademark/trade dress tied to product and marketing; patents generally excluded.
    - privacy & network_security: Third-party liability from breaches/security failures; table-stakes for data handlers.
    - business_interruption: I record system failure (non-malicious) and dependent providers; cloud reliance makes this a key differentiator.
    - regulatory: Defense/penalties where insurable—AG/FTC/DPAs are realistic for growth startups.
    - pci: Only if you touch card data, but it's commonly assumed—captured explicitly.
    - breach_response: Legal, forensics, notification, PR—your first 72 hours; panel access matters.
    - cyber_extortion: Ransomware costs—frequent loss driver.
    - social_engineering: Often optional/excluded; real treasury risk.
    - contractual_liability_carveback: Preserves liability that exists absent the contract and other practical carve-backs negotiated in enterprise deals.
- exclusions: Where expectations get reset. I list major exclusions (e.g., patents, BI/PD, utilities/infrastructure outside control, prior acts/notice, insured-vs-insured, TCPA/CAN-SPAM) with any carve-backs so the gaps are explicit.
- definitions: Precision prevents disputes. I capture professional_services, technology_services, and wrongful_act texts because they bound E&O/media scope and are frequent declination levers when vague.
- conditions: Operational essentials—notice/reporting mechanics, ERP/tail, territory/jurisdiction—so legal/ops can build compliant workflows and plan for M&A/tail needs.
- unknowns: Transparent gaps when Declarations/endorsements aren't public, with search trails to close via broker quotes, binders, or SERFF.

Scoring methodology and rationale

- What is scored: Each policy's `output.yaml` is evaluated on 0–5 sub-factors grouped as:
  - core: trigger, defense_costs, consent_settle_hammer, erp_tail, retro_date
  - coverage: tech_prof_services, tech_products, media_ip, privacy_regulatory, business_interruption_dbi, pci_reputational
  - defs_carves: defs_clarity, carvebacks_major_exclusions
  - limits: sublimits_breach_bi_reg_pci, retentions_alignment
  - ops: panel_counsel_flex, notice_practicality_services
- Normalization: Per group g, scores are clamped to [0,5] and scaled to [0,1] as $n_g = \frac{\sum_i \min(5,\max(0,s_i))}{5 \cdot N_g}$.
- Aggregation: Profile-weighted sum, scaled to 0–100, plus additive adjustments: $\text{Score} = 100 \cdot \sum_g w_g \, n_g + \text{bonus\_penalties}$.
- Weights by profile:
  - default: core 0.30, coverage 0.30, defs_carves 0.20, limits 0.10, ops 0.10
  - b2b_saas: core 0.25, coverage 0.35, defs_carves 0.25, limits 0.10, ops 0.05
  - consumer: core 0.25, coverage 0.35, defs_carves 0.20, limits 0.10, ops 0.10
- Bonuses/penalties (additive):
  - Defense costs inside limits: −10 if per-claim limit ≤ $2M, else −5
  - Hammer: hard_100 −10; soft_50 −3; none_or_consent_only +3
  - Incident response panel: +2
  - Dependent BI endorsement: +3
- Scenarios: Two checks are reported (SaaS API outage/SLA; UGC defamation takedown) to aid interpretation; they do not change totals.
- Ranking: Policies are sorted by the chosen profile score; top 4 are flagged.
- Rationale: Emphasize coverage breadth and foundational terms (core) first; clarity of definitions/carve-backs next (reduces denial/coverage disputes). Limits/ops influence practical fit but less the base coverage, so lower weights. B2B SaaS places extra weight on coverage and carve-backs (contracts/SLA/BI); consumer weighting restores ops for notice/panel usability. Penalties address limit erosion and settlement friction; bonuses reward practical first-response and dependency coverage common to startups.

Key results (most important)

- Catalog coverage: 19 providers across major carriers (e.g., Beazley, Chubb, AXA XL, CNA, AIG, Zurich, Hartford, Hanover, Markel) and startup-focused MGAs (Coalition, At-Bay, Corvus, Vouch, Embroker).
- Common structure observed:

    - Claims-made (often claims-made-and-reported), duty-to-defend, defense costs typically inside limits.
    - Consent-to-settle with hammer clauses (soft hammers most common; % varies by form).
    - Business interruption increasingly split by trigger (malicious act vs system failure) and dependent providers; several forms restrict BI to insured's own network.
    - Contractual liability exclusions with targeted carve-backs (liability absent contract, privacy policy breach, PCI, and/or unintentional breach of a written services contract).
    - Broad privacy/security and media definitions; patents generally excluded; trade secret carve-backs vary.
- Selected 4 (plus CFC) — highlights:
    - Corvus Smart Tech E&O: Strong pairing of Tech/Professional Services liability with cyber; dynamic loss-prevention services; practical carve-backs (incl. privacy-linked trade secret); well-balanced third- and first-party modules.
    - Embroker (with Everspan): Duty-to-defend; modern sublimits (e.g., reputational harm, bricking, betterment); broad regulatory and contract carve-backs; good fit for venture-backed growth companies.
    - Beazley MediaTech: Mature integrated form (Tech E&O + Media + Cyber); clear definitions and contractual carve-backs; robust first-party suite; widely recognized market option.
    - AXA XL CyberRiskConnect: Unified coverage with solid privacy/security and tech services liability; BI tends to focus on own network; exclusions/limitations are explicit and verifiable; strong enterprise credibility.
    - CFC (reference): Included for founder comparability; widely used in startup market; adds context across cyber/E&O modules (specimen provided separately by Corgi).

Notes

- Assumption: No predefined selection criteria were provided; I analyzed broadly, then selected 4 policies that a startup founder can compare quickly and use to decide, balancing coverage breadth, clarity of carve-backs, and verifiability of terms.