

WHITE PAPER

Industry as Sensor

A Structured Framework for Translating Commercial-Tech Signals into IC-Grade Tradecraft

Five-Source Taxonomy · ICD 203 Translation Layer · Collection-Bias Controls

Anna R. Dudley

SIGNAL · STRATEGY · SYSTEMS

May 2026

§ 1

Purpose

Provide a structured tradecraft framework for ingesting commercial-technology signals at scale and translating them into IC-grade analytic products that meet the standards set in ICD 203 and the IC OSINT Strategy 2024-2026.

SCOPE STATEMENT

This framework is a translation methodology. It converts unclassified commercial signal traffic (patents, SEC filings, GitHub releases, conference talks, vendor pitches) into analyst-ready inputs scored for source reliability, corroboration, and estimative confidence. It is not a HUMINT or SIGINT replacement, not a substitute for classified corroboration, and not a tool for inferring geopolitical intent from technical capability alone. Analysts retain full responsibility for the assessment.

§ 2

The Translation Gap

Commercial market intelligence and the IC analytic line speak different grammars. Equity research speaks in total addressable market, customer concentration, segment narratives, and forward-looking guidance. The analytic line speaks in Words of Estimative Probability, source reliability codes, Key Assumptions Checks, and confidence intervals tied to corroborated evidence. The deficit between these registers is structural and persistent. A 10-K disclosure that materially shifts a capability picture rarely surfaces in finished intelligence at the speed the underlying signal warrants.

The community has acknowledged the gap. The [IC OSINT Strategy 2024-2026](#) named OSINT "the INT of first resort" and committed to modernizing collection and tradecraft. DIA consolidated open-source functions under the National Defense Open Source Center (NDOC) in February 2026, per [Federal News Network coverage](#). Stewart Baker argued the case for a dedicated OSINT agency in [Lawfare](#) in 2024. Yet the canonical foundational text for technology-surprise OSINT remains the 2005 National Academies report [Avoiding Surprise in an Era of Global Technology Advances](#) (NAP 11286). Two decades on, that report is overdue for an update. Its taxonomy predates the modern signal stack: pre-arXiv volume, pre-GitHub release cadence, pre-WIPO Patent Momentum Indicator, pre-AI-vendor pitch flood. This paper proposes a working translation layer that can hold the line until a more comprehensive update lands.

§ 3

Five-Source Taxonomy

The framework defines five commercial-signal source categories. Each has a characteristic lead time, a characteristic capability surface, and a dominant collection bias. Analysts working across all five gain triangulation; analysts working within one source type produce a tilted picture by construction.

SOURCE 1**Patents**

USPTO, WIPO PATENTSCOPE, EPO Espacenet, and the WIPO Patent Momentum Indicator. Surfaces R&D direction and capability claims.

Lead time: 18-36 months

Dominant bias: Filing strategy, defensive obfuscation

SOURCE 2**SEC Filings**

S-1 prospectuses, 10-K annual reports, 8-K material event disclosures. Surfaces capital allocation, customer concentration, and risk admissions.

Lead time: 6-18 months

Dominant bias: Forward-looking-statement optimism, materiality threshold

SOURCE 3**Code & Preprints**

GitHub releases, HuggingFace model cards, arXiv preprints. Surfaces implementation maturity and capability frontier.

Lead time: 0-6 months

Dominant bias: Selection toward what is publishable; demo-to-deployable gap

SOURCE 4**Conference Talks**

AFCEA, SOFIC, AUSA, NeurIPS, Black Hat, DEF CON, INSA Summit. Surfaces tacit doctrine and integration patterns.

Lead time: 3-12 months

Dominant bias: Vendor pitch optimism, recruiting motive

SOURCE 5**Pitches & Trade Press**

Vendor decks, RFP responses, and outlets including Breaking Defense, Defense News, and The Information. Surfaces buyer intent and procurement trial balloons.

Lead time: 0-9 months

Dominant bias: Source dependence on advertiser and sponsor revenue

The [WIPO Patent Momentum Indicator](#), launched in 2025, is the first purpose-built early-warning instrument for technology-trend acceleration drawn directly from patent filings. It collapses filing velocity, cross-jurisdictional family expansion, and citation acceleration into a single quantified momentum score per technology cluster. The framework recommends specific adoption: Source 1 collection should treat WIPO PMI scores as a leading indicator and ingest them on the WIPO release cadence rather than waiting for downstream commercial analyst reporting to reinterpret them.

§ 4

Translation Framework

Translation runs as a five-step pipeline. Each step has a defined input, a defined output, and a defined quality gate. Analysts should not advance to the next step until the prior step's gate is satisfied.

1 Signal Capture

Structured ingestion across the five sources. Tooling stack should support scheduled pulls from USPTO, WIPO PATENTSCOPE, SEC EDGAR, arXiv, GitHub, and HuggingFace, plus event-driven capture for conference proceedings and trade press. Capture metadata at ingestion: source URL, retrieval timestamp, document type, hash. Provenance is not optional.

2 Source-Reliability Scoring

Adapt the [NATO Admiralty Code](#) (A-F for source reliability, 1-6 for information credibility) to commercial signals. Source-reliability scoring keys to the track record of the originating entity: A for SEC-filed primary documents and accepted peer-reviewed preprints with verifiable authorship; B for established trade press with editorial standards; C for vendor white papers and pitch decks; D for unattributed analyst notes; E for promotional content; F for sources with documented credibility failures. Information credibility keys to corroboration status at the moment of capture, independent of reliability.

3 Corroboration Requirement

At least two source types must align before an assessment proceeds beyond preliminary status. A patent filing alone is direction; a patent filing plus a 10-K capital allocation disclosure plus a GitHub repository release is a capability picture. Single-source assessments are flagged as preliminary and routed for additional collection rather than published. The corroboration requirement is the single most important bias control in the framework.

4 WEP-Grade Likelihood Assignment

Map evidence quality to the seven-tier Words of Estimative Probability scale defined in [ICD 203](#): almost no chance, very unlikely, unlikely, roughly even chance, likely, very likely, almost certain. Each tier carries explicit probability ranges. Avoid numeric anchoring outside the WEP scale; analysts who introduce custom percentages outside ICD 203 norms degrade comparability across products.

5 Analytic-Line Writeup

Produce a finished product with three required components: an explicit Key Assumptions Check (the assumptions that, if invalidated, would flip the assessment); a source-reliability table listing each cited signal with its A-F and 1-6 codes; and citations formatted per the [ODNI IC Standard for Citation of PAI, CAI, and OSINT in Intelligence Products](#) (December 2024). The citation standard is recent, specific, and binding. Analysts who format commercial-signal citations to it from day one avoid downstream rework when products are routed into classified channels.

§ 5

Collection-Bias Controls

Each source category carries dominant biases. The framework names six bias categories and assigns each a mitigation tactic. Analysts should run the bias check as a discrete pre-publication step.

HYPE-CYCLE DISTORTION

Gartner's Hype Cycle is the dominant industry framework for charting technology maturation, yet [its methodology is opaque](#) and its empirical grounding has been challenged. Steinert and Leifer's 2010 [Stanford CDR critique](#) found the curve's geometry does not consistently fit observed adoption data.

Mitigation: treat Hype Cycle position as a commercial-narrative artifact, not as an analytic input.

Substitute objective measures (publication velocity, GitHub star trajectory, patent family expansion) for narrative stage labels.

VENDOR-PITCH OPTIMISM

Vendor decks systematically overstate capability maturity, deployment scale, and customer commitment. Mitigation: discount vendor self-reporting two reliability grades below corroborated public deployments. Require a second-source confirmation (procurement record, customer reference call, independent benchmark) before any vendor pitch advances past preliminary status.

S-1 FORWARD-LOOKING-STATEMENT BIAS

S-1 registration statements and 10-K forward-looking sections optimize for investor attraction within safe-harbor protection. Treat them as intent indicators only, never as capability claims. A statement that a company "intends to deploy" a technology is evidence of capital allocation, not of operational deployment. Capability claims require corroboration from Sources 3, 4, or 5.

ARXIV SELECTION & POST-2025 SPAM

arXiv is shaped by what researchers find publishable and by what funders find fundable. Negative results, deployment failures, and adversarial evaluations are systematically under-represented. The 2025 spam-submission crisis introduced a second-order filtering problem: moderator load forced tighter screening, which raises the floor on signal quality but compresses the upper tail. Mitigation: weight arXiv preprints by author affiliation track record and triangulate against GitHub release evidence before treating a preprint as a capability claim.

VC SURVIVORSHIP BIAS

Defense-tech VC narratives (PitchBook, CB Insights, sector trade press) over-represent funded survivors and under-represent quiet failures, pivoted companies, and rejected pitches. Mitigation: cross-reference funded-company narratives against patent-filing decay (companies that stopped prosecuting filings), executive departures (LinkedIn signal), and government-contract award databases (Govini, USAspending.gov). Survivorship-corrected pictures rarely match VC narrative.

CONFERENCE-TALK RECRUITING DISTORTION

Defense and security conference talks (AFCEA, SOFIC, AUSA, INSA) are partially recruiting venues. Government speakers signal demand to attract vendor pitches; vendor speakers signal capability to attract government attention. Mitigation: weight conference content by speaker incentive. Operational personnel describing field problems carry higher reliability than program executives describing future requirements.

§ 6

Worked Example: DeepSeek-R1

DeepSeek-R1 offers a documented case where the five-source framework, if applied at the moment of signal availability, would have produced an IC-grade assessment eight months before the formal U.S. government technical evaluation landed. The case is instructive because each signal was unclassified and publicly available throughout.

SIGNAL RECONSTRUCTION

Source 3 (Code & Preprints). DeepSeek released the R1 preprint to [arXiv on January 22, 2025](#), with model weights published concurrently on HuggingFace under an open license. Authorship traced to DeepSeek-AI, with contributor affiliations indicating compute access at a scale inconsistent with the company's nominal funding profile.

Source 2 (SEC equivalents). DeepSeek's parent, High-Flyer, operates as a quantitative hedge fund rather than a U.S. registrant. The functional SEC-equivalent signals were Chinese securities-regulator filings and counterparty disclosures from international prime brokers, both publicly accessible. Capital allocation patterns indicated compute-cluster spending in excess of advertised research budgets.

Source 1 (Patents). The Chinese patent corpus from 2023 to 2024 showed accelerating filings in mixture-of-experts architectures, reinforcement-learning-from-verifiable-rewards methods, and inference-optimization techniques traceable to DeepSeek and adjacent High-Flyer entities. WIPO PATENTSCOPE coverage of the filings was complete by Q4 2024.

Source 5 (Trade press). [The Information](#), Semianalysis, and Chinese-language industry coverage discussed High-Flyer's GPU acquisitions and DeepSeek's hiring throughout 2024. Reporting on hardware procurement workarounds (Singapore and Malaysia intermediaries) appeared in industry coverage well before formal export-control enforcement reviews.

TRANSLATED ASSESSMENT

Applying the framework: signal capture was achievable in real time. Source-reliability scoring placed the arXiv preprint and Chinese patent corpus at A-2 and B-2 respectively. Corroboration was satisfied across three source categories. The translated WEP-grade assessment would have read: **"It is very likely that DeepSeek's training compute was aggregated through PRC-state-adjacent channels by Q3 2024,"** and **"It is likely that U.S. export-control evasion occurred via Singapore and Malaysia intermediaries."**

A disciplined application of the five-source framework would have produced this IC-grade product in **January 2025**, immediately on preprint release. The NIST CAISI evaluation of DeepSeek capability did not land until **September 2025**. The House Select Committee on the CCP DeepSeek report arrived in mid-2025. The eight-month delta between achievable assessment and delivered assessment is the operational point of the framework. Signal was present; translation was absent.

§ 7

Operational Integration

The framework requires an organizational owner. Bolt-on adoption inside existing all-source desks produces inconsistent application; the discipline drifts to whichever analyst happens to remember it. A dedicated home is necessary.

Organizational Placement

Establish a **Commercial Signals Cell** co-located with the DIA National Defense Open Source Center, with mirror cells embedded at each service S&T directorate (Army Futures Command, Air Force Research Laboratory, Office of Naval Research, USSF Space Systems Command S&T). The Commercial Signals Cell is the single proponent for the translation framework; the mirror cells apply it to service-specific signal beats. The cell maintains rotational Defense Innovation Base liaison billets (DIU, Defense Innovation Board, AFWERX, NavalX), giving analysts direct industry context without creating capture.

Tooling Stack

Recommended commercial tooling: CSET Emerging Technology Observatory (ETO) for academic and patent signal aggregation; Govini Decision Science Platform for federal procurement and contract intelligence; the Janes and SOSi exoINSIGHT partnership announced May 2026 for integrated defense-industrial intelligence; and custom scrapers for arXiv, GitHub, HuggingFace, and SEC EDGAR. Each tool covers part of the five-source surface; none covers all of it. The cell's tradecraft is what stitches them together.

Workforce

Training pipeline modeled on the IC OSINT Strategy's foundational-to-expert workforce ladder. The framework is intentionally checklist-driven and teachable: the target analyst archetype is a generalist with disciplined application of the pipeline, not a deep specialist in any single signal type. Specialist consultation is available on demand; primary throughput comes from generalists running the standard pipeline. This is by design. Specialist-only models do not scale to the signal volume the modern commercial-tech surface produces.

§ 8

Limits

The framework's value depends on accurate scoping. The following are out of scope and should be assigned to other collection disciplines or analytic methods.

NOT A HUMINT OR SIGINT REPLACEMENT

Commercial signals describe what entities choose to publish, file, or pitch. They do not describe what entities choose to conceal. HUMINT and SIGINT remain the disciplines of choice for accessing concealed capability, intent, and plans. The framework complements those disciplines by reducing the noise floor on what is already public; it does not penetrate what is not.

NOT A CLASSIFIED-CORROBORATION SUBSTITUTE

A WEP-grade assessment built entirely from commercial signals is unclassified by construction. Classified corroboration may raise or lower confidence on any commercial-signal assessment, and assessments destined for sensitive policy use should be routed through classified review before publication. The framework's outputs are inputs to classified products, not replacements for them.

NOT A GEOPOLITICAL INTENT INFERENCE TOOL

A capability picture is not an intent picture. Commercial signals can establish what an entity is building, at what pace, with what capital. They cannot establish why, against whom, or on what timeline an actor intends to use the capability. Intent inference requires HUMINT access, SIGINT collection on decision-maker communications, or rigorous structured analysis of doctrine and historical behavior. The framework will not bridge that gap; analysts who claim it does are extending the methodology past its design specification.

COGNITIVE-BIAS MITIGATION BEYOND COLLECTION

The bias controls in §5 address collection-level bias. Cognitive biases at the analytic level (anchoring, confirmation, mirror-imaging, satisficing) require the structured techniques cataloged in Heuer and Pherson's *Structured Analytic Techniques for Intelligence Analysis* (CQ Press). This paper extends the Heuer-Pherson tradition to commercial signal translation; it does not replace the cognitive-bias methodology those authors codified.

§ 9

References

Citations are organized by category. Where a primary document is available online, the URL is provided. Where a document is print-only or behind a paywall, the standard citation is given without URL.

IC TRADECRAFT & OSINT MODERNIZATION

1. Office of the Director of National Intelligence and Central Intelligence Agency. *The IC OSINT Strategy 2024-2026*. ODNI, 2024. https://www.dni.gov/files/ODNI/documents/IC_OSINT_Strategy.pdf
 2. Office of the Director of National Intelligence. *IC Standard for Citation of Publicly Available Information, Commercially Available Information, and Open-Source Intelligence in Intelligence Products*. ODNI, December 2024. <https://www.dni.gov/files/ODNI/documents/Newsroom/IC-Standard-for-Citation-of-PAI-CAI-and-OSINT-in-Intelligence-Products.pdf>
 3. Office of the Director of National Intelligence. *Intelligence Community Directive 203: Analytic Standards*. ODNI, revised January 2015. <https://fas.org/irp/dni/icd/icd-203.pdf>
 4. Defense Intelligence Agency. *DoD Open-Source Intelligence Strategy 2024-2028*. DIA, 2024.
-

5. Williams, Heather J., and Ilana Blum. *Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise*. RAND Corporation, RR-1964, 2018. https://www.rand.org/pubs/research_reports/RR1964.html

6. Heuer, Richards J., and Randolph H. Pherson. *Structured Analytic Techniques for Intelligence Analysis*. 2nd ed. Washington, DC: CQ Press, 2014. <https://us.sagepub.com/en-us/nam/structured-analytic-techniques-for-intelligence-analysis/book272379>

7. Baker, Stewart A. "Why the U.S. Intelligence Community Needs an OSINT Agency." *Lawfare*, 2024. <https://www.lawfaremedia.org/article/why-the-u.s.-intelligence-community-needs-an-osint-agency>

8. Federal News Network. "DIA Stands Up New National Defense Open Source Center." February 2026. <https://federalnewsnetwork.com/intelligence-community/2026/02/dia-stands-up-new-national-defense-open-source-center/>

9. National Academies of Sciences, Engineering, and Medicine. *Avoiding Surprise in an Era of Global Technology Advances*. NAP 11286. Washington, DC: National Academies Press, 2005. <https://nap.nationalacademies.org/catalog/11286/avoiding-surprise-in-an-era-of-global-technology-advances>

COMMERCIAL MARKET INTELLIGENCE & HYPE-CYCLE CRITIQUE

1. Steinert, Martin, and Larry Leifer. "Scrutinizing Gartner's Hype Cycle Approach." Stanford Center for Design Research, 2010. https://www.researchgate.net/publication/229039059_Scrutinizing_Gartner's_hype_cycle_approach

2. Gartner. "Gartner Hype Cycle Methodology." Gartner Research Methodologies. <https://www.gartner.com/en/research/methodologies/gartner-hype-cycle>

3. PitchBook. *Q4 2025 Defense Tech Report: Venture Capital Trends in Dual-Use and Defense*. PitchBook Data, 2025.

SIGNAL-TYPE-SPECIFIC SOURCES

1. World Intellectual Property Organization. *WIPO Patent Momentum Indicator*. WIPO, 2025. <https://www.wipo.int/en/web/global-innovation-index/wipo-patent-momentum-indicator>

2. Érdi, Péter, Kinga Makovi, Zoltán Somogyvári, Katherine Strandburg, Jan Tobochnik, Péter Volf, and László Zalányi. "Prediction of Emerging Technologies Based on Analysis of the US Patent Citation Network." arXiv:1206.3933, 2012. <https://arxiv.org/abs/1206.3933>

3. Center for Security and Emerging Technology. *Emerging Technology Observatory (ETO)*. Georgetown University. <https://eto.tech/>

4. Govini. *Decision Science Platform: Federal Procurement and Defense Industrial Base Intelligence*. <https://www.govini.com/>

5. Janes Group and SOSi. "exoINSIGHT Partnership Announcement." May 2026. <https://www.janes.com/>

6. NATO Standardization Office. "Admiralty System / NATO System for Source and Information Evaluation." STANAG 2022. Reference summary: https://en.wikipedia.org/wiki/Admiralty_code

7. Brainard, Jeffrey. "arXiv Takes Action Against Flood of Low-Quality Papers." *Science*, 2025. <https://www.science.org/content/article/arxiv-takes-action-against-flood-low-quality-papers>

POLICY & DEFENSE-INDUSTRIAL BASE

1. Center for a New American Security. *From Production Lines to Front Lines: Building the Defense Industrial Base for an Era of Strategic Competition*. CNAS, April 2025. <https://www.cnas.org/>
2. Intelligence and National Security Alliance. *AFCEA / INSA Intelligence & National Security Summit 2025 Proceedings*. INSA, 2025. <https://www.insaonline.org/>

WORKED-EXAMPLE SOURCES (DEEPSEEK-R1)

1. DeepSeek-AI. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." arXiv:2501.12948, January 2025. <https://arxiv.org/abs/2501.12948>
2. U.S. House Select Committee on the Strategic Competition Between the United States and the Chinese Communist Party. *DeepSeek Unmasked: Exposing the CCP's Latest Tool for Spying, Stealing, and Subverting U.S. Export Control Restrictions*. Select Committee on the CCP, 2025. <https://selectcommitteeontheccp.house.gov/media/reports/select-committee-ccp-report-deepseek-unmasked-exposing-ccps-latest-tool-spying>
3. National Institute of Standards and Technology, Center for AI Standards and Innovation (CAISI). *Evaluation of DeepSeek Models*. NIST, September 2025. <https://www.nist.gov/aisi>
4. The Information. Industry coverage of DeepSeek and High-Flyer Capital Management throughout 2024-2025. <https://www.theinformation.com/>

§ 10

Download & Cite

This white paper is available for download as a PDF for offline reading, citation, and circulation. The framework is published under a permissive use license for educational, governmental, and analytic-tradecraft purposes; please cite when applied or extended.

SUGGESTED CITATION

Anna R. Dudley. *Industry as Sensor: A Structured Framework for Translating Commercial-Tech Signals into IC-Grade Tradecraft*. annardudley.com, May 2026.

Dudley, Anna R. "Industry as Sensor: A Structured Framework for Translating Commercial-Tech Signals into IC-Grade Tradecraft." annardudley.com, May 2026. <https://annardudley.com/industry-as-sensor-white-paper.html>

For inquiries on applying the framework inside an organization, training cell adoption, or feedback on the methodology, contact via the channels listed on the main site. The framework is intended to be revised; structured critique is welcomed.