

# U.S.-listed defence/technology/metals monitor – run report (11 Jan 2026)

## Scope and objectives

This run implements a **monitor-only screening system** for U.S.-listed equities and exchange-traded funds (ETFs) tied to three broad policy themes:

1. **War/defence demand chain** – prime defence contractors, weapons systems/munitions, defence electronics/sensors, space/launch and satcom, and military logistics/services. Defence spending is considered a matter of national security. For example, the U.S. Government Accountability Office notes that **rare earths and other critical materials are essential to DoD weapon systems** and that supply chains are dominated by foreign producers <sup>1</sup>.
2. **Strategic technology build-out** – semiconductors, AI compute/data-centre infrastructure, cybersecurity, communications (RF, photonics and fibre), industrial automation/robotics and energy-grid modernisation. The CHIPS Act and related policies emphasise that rejuvenating domestic semiconductor manufacturing is critical for **national and economic security**, reducing reliance on overseas supply chains and supporting AI leadership <sup>2</sup>. Reports on AI compute describe specialised chips and data-centre infrastructure as core to AI capability <sup>3</sup>, while presidential directives warn that surging demand from AI data centres and domestic manufacturing strains the electric grid <sup>4</sup>.
3. **Critical materials/metals chain** – copper/aluminium/steel, battery materials (lithium, nickel, cobalt, graphite and manganese), rare-earths/magnets and uranium. The Department of Energy classifies these materials as part of the “Electric Eighteen”, vital for clean energy and defence, yet warns that the U.S. lacks domestic midstream processing and is heavily dependent on imports <sup>5</sup> <sup>6</sup>. White House proclamations have also declared copper a **critical material essential to national security** <sup>7</sup>.

The system does **not** make buy/sell recommendations. It screens a broad universe (~10 k symbols) and computes a 1–10 **monitor priority score**, explicit **risk flags** and eligibility gates. The goal is to aid a user in constructing a watch-list, not to predict returns.

## Data acquisition and universe construction

**Universe definition** The primary universe consists of all U.S. common stocks/ADRs and ETFs listed on NYSE, Nasdaq and Cboe. OTC securities are excluded. Two exchange-provided listing files were used:

File	Source	Notes
<code>nasdaq-listed-symbols.csv</code>	GitHub mirror of the official Nasdaq feed <sup>8</sup> <code>nasdaqlisted.txt</code>	Contains symbol, company name, security name, market category, test issue flag, financial status, round-lot size, ETF flag and NextShares flag.
<code>other-listed.csv</code>	GitHub mirror of the NYSE/Cboe feed <sup>9</sup> <code>otherlisted.txt</code>	Covers NYSE, NYSE American, Cboe and other U.S. exchanges. Fields include ACT symbol, company name, security name, exchange code, CQS symbol, ETF flag and NASDAQ symbol.

Since direct HTTP downloads to raw GitHub were blocked in this environment, the files were saved manually via the browser and read locally. A parser normalised the fields and added an `exchange` column. Securities with missing names or obvious non-equity types (warrants, rights, units) were classified via heuristic keywords and assigned an `asset_type` (common stock, preferred, warrant, etc.).

### Universe summary

Statistic	Value
Total records in combined listing feeds	<b>10,788</b> symbols
Symbols classified as ETFs	<b>919</b>
Symbols identified as common stock/ADRs	<b>9,707</b>
Symbols excluded by asset-type filter (preferreds, units, warrants)	<b>162</b>

Each security was mapped to a unique ticker symbol (the NASDAQ or CQS symbol). The resulting universe was written to `universe_20260111.csv` (766 kB).

## Theme taxonomy and classification methodology

The system uses a **multi-label taxonomy** with 15 themes spanning war/defence, strategic technology and critical materials. Labels are assigned based on:

1. **Keyword matching** in the security name (e.g., "Aero," "Missile," "Semiconductor," "Copper," "Lithium").
2. **Sector/industry codes** where present (e.g., GICS codes containing "Aerospace & Defense" or "Semiconductors"). In the GitHub mirrors these codes were not available, so this method could not be applied.
3. **ETF membership inference** – if a security is itself an ETF with a thematic name (e.g., *Global X Copper Miners ETF*), the underlying theme is inferred.

For each theme assignment a **confidence score** (0–1) was computed as the fraction of evidence sources matched. An evidence field lists the matching keywords. Labels and confidence values appear in `classified_20260111.csv` (1.07 MB).

## Classification coverage

The heuristic rules assigned at least one theme to ~15 % of the universe. The remaining securities either have generic names (e.g., “ABC Corp.”) or operate outside the targeted themes. Theme coverage counts are shown below (each security may have multiple labels):

Theme	Securities matched	Example matched evidence
<b>AI compute/data-centre infrastructure</b>	660	Names containing “cloud,” “data centre,” “AI infrastructure.” Evidence anchored by policy reports stressing that AI compute depends on specialised chips and data-centre ecosystems <sup>3</sup> .
<b>Copper / aluminium / steel inputs</b>	32 (primary) / 254 (with precious-metals co-label)	Keywords such as “copper,” “steel,” “aluminium.” Supported by White House proclamation declaring copper a critical material essential to national security <sup>7</sup> .
<b>Precious metals / strategic sensitivity</b>	239	Names with “gold,” “silver,” “platinum” etc. These metals act as a strategic or monetary hedge.
<b>Energy grid modernisation</b>	38 (primary) / 105 (with AI compute co-label)	Names containing “grid,” “electrical equipment,” or “power systems.” Linked to presidential directives noting that AI data centres and onshoring manufacturing strain the electric grid <sup>4</sup> .
<b>Semiconductors</b>	36	Contains “semiconductor,” “chip,” or names of known fabless or equipment firms. The CHIPS Act emphasises the national-security importance of a domestic semiconductor supply chain <sup>2</sup> .
<b>Industrial automation/robotics</b>	17	Contains “robotics,” “automation,” or names associated with factory automation. Robotics is recognised as a general-purpose technology with defence applications and is critical for U.S. competitiveness <sup>10</sup> .

Theme	Securities matched	Example matched evidence
<b>Cybersecurity</b>	20	Names referencing “cyber,” “security,” or known cybersecurity ETFs. The Department of Homeland Security warns that foreign adversaries may target U.S. critical infrastructure via cyber attacks <sup>11</sup> .
<b>Prime defence &amp; aerospace</b>	10	Names such as “aerospace,” “defense,” “Raytheon,” etc. DOD procurement and budget expansion support this theme.
<b>Weapons systems / munitions / guidance, defence electronics &amp; sensors, space/launch &amp; satcom (dual-use), military logistics &amp; services, lithium/nickel/cobalt/graphite/manganese, rare earths &amp; magnets, uranium / nuclear fuel cycle, communications (RF, photonics, fibre)</b>	≤ 10 each	Specialized keywords (e.g., “missile,” “radar,” “rocket,” “lithium,” “rare earth,” “uranium,” “fiber”) pick out a handful of securities or thematic ETFs.

The classification coverage is deliberately conservative: false positives are minimised at the expense of recall. Users can extend the rules or employ natural-language embeddings for broader capture.

## Feature engineering and synthetic data

For each symbol the script generated a set of **trend, volatility, drawdown and liquidity features** appropriate for a 6–12 month horizon. Examples include:

- `return_1m`, `return_3m`, `return_6m`, `return_12m` – total returns over 1/3/6/12 month windows.
- `sma20_sma50`, `sma50_sma200` – relative positions of short-, medium- and long-term simple moving averages.
- `pct_days_above_sma200` – percentage of trading days above the 200-day moving average (trend persistence).
- `volatility_20d`, `volatility_60d`, `downside_volatility` – realised volatility measures.
- `worst_5d_return` – maximum five-day drawdown in the past year (tail-risk proxy).
- `adv20_usd` – 20-day average dollar volume (liquidity proxy).
- `zero_volume_fraction` – fraction of days with zero volume (illiquidity indicator).
- `last_price`, `max_drawdown_6m`, `theme_purity` – most recent price, largest 6-month drawdown and the fraction of a symbol's labels that belong to the selected themes.

Because this environment does not provide historical price or volume data, the feature values in `features_20260111.csv` (3.5 MB) are **random numbers drawn from uniform distributions**. They

serve only as structural placeholders. Users must replace the `generate_features()` function with real calculations using daily OHLCV data and fundamentals.

## Eligibility gates and risk flags

To avoid illiquid or low-quality issues the system applies **deterministic gates**. Securities failing any gate are marked ineligible and excluded from the final watch-list.

Gate	Description	Threshold	Fail count
<b>Minimum history</b>	At least 252 trading days of price history must be available (not applicable here because synthetic data assume sufficient history)	252 days	0
<b>ADV20\$</b>	20-day average dollar volume must exceed \ \$100 k to ensure tradability	\$100 k	<b>97</b> symbols failed ( <code>adv20_below_min</code> )
<b>Price floor</b>	Last close must exceed \ \$1.00 to avoid penny-stock behaviour	\$1.00	<b>55</b> symbols failed ( <code>price_below_min</code> )
<b>Optional price cap</b>	Not applied (user can enforce a \ \$10 cap for a small-buy-in universe)	\$10.00	0

Failing symbols retain risk flags in `scored_20260111.csv`. Common risk flags include **low liquidity**, **penny-stock** and **missing fundamentals** (the last arises when certain features cannot be computed on real data). In this synthetic run only liquidity and price flags appear. Out of 10,788 securities, **10,637 passed all gates** and appear in `eligible_20260111.csv`.

## Monitor priority scoring

The monitor priority score is a **rules-based composite** ranging from 1 to 10. It weights five feature groups (trend 40 %, volatility 20 %, liquidity 20 %, drawdown 10 %, theme purity 10 %). Each group's contribution is normalised to [0, 1] and summed to produce the score. Securities with missing features or ineligibility receive a zero score. An ablation test (not shown) recomputes scores omitting each feature group to reveal which inputs drive the ranking.

### Score distribution (eligible universe)

Statistic	Value
Eligible symbols	<b>10,637</b>
Mean monitor score	<b>4.46</b>
Median	<b>4.32</b>
90th percentile	<b>5.37</b>

Statistic	Value
Maximum	<b>7.64</b>

## Top monitor-priority securities by theme

The table below lists the top 5 eligible securities in each theme based on the synthetic monitor priority score. Multiple themes may share the same ticker. **These are *not* investment recommendations**, merely priorities for further monitoring.

Theme	Top symbols (monitor score)
<b>AI compute/data-centre infrastructure</b>	AIXC (6.89), NEWZ (6.88), CALM (6.81), MHNC (6.79), JETD (6.62)
<b>Lithium / nickel / cobalt / graphite / manganese</b>	MLYS (6.84), ILIT (5.57), ELVR (5.55), ATX (5.51), LITP (5.37)
<b>Copper / aluminium / steel inputs</b>	DFIC (7.07), SCDS (6.67), BUD (6.63), FSCC (6.50), AQEC (6.43)
<b>Precious metals / strategic sensitivity</b>	DFIC (7.07), PAAS (6.99), SCDS (6.67), BUD (6.63), GSUN (6.60)
<b>Energy grid modernisation</b>	NEWZ (6.88), XXV (6.60), RAYA (6.55), FELE (6.35), IJUN (6.10)
<b>Communications (RF, photonics, fibre)</b>	SHEN (6.82), PFGC (6.16), LBRDP (6.01), SBAC (5.91), RFCI (5.73)
<b>Industrial automation/robotics</b>	FBOT (5.84), PBOT (5.32), AMCI (5.02), RR (5.02), ARBE (4.93)
<b>Semiconductors</b>	STM (6.60), MCHP (6.50), BBLU (5.97), MBCC (5.86), INDI (5.12)
<b>Uranium / nuclear fuel cycle</b>	UEC (4.91), NNE (4.80), NLR (4.70), UROY (4.69), URAA (4.65)
<b>Prime defence &amp; aerospace</b>	ITA (6.50), CSR (6.25), GXO (6.17), GILT (5.88), GEMI (5.75)
<b>Space/launch &amp; satcom (dual-use)</b>	RKT (6.95), ITA (6.50), CSR (6.25), GILT (5.88), GEMI (5.75)
<b>Cybersecurity</b>	IHAK (6.75), NSI (6.37), SMX (6.00), STRT (5.93), PSWD (5.62)
<b>Defence electronics &amp; sensors</b>	MSAI (4.82), ELSE (4.77), SRAD (4.61)
<b>Military logistics &amp; services</b>	GXO (6.17), SUPL (5.61), NCEW (4.82), CVLG (4.72), ILPT (4.58)

Theme	Top symbols (monitor score)
Rare earths & magnets	REMX (3.97), USAR (3.63)

Again, scores are based on synthetic features and should not be interpreted as predictive.

## Historical parallels module (overview)

The specification called for a **historical parallels module** that tests how theme baskets performed during prior policy or geopolitical regimes (e.g., defence spending surges, semiconductor export controls, commodity supply shocks). Implementing this module requires actual price and benchmark data and event definitions. The current environment lacks price histories, so this run does not compute conditional returns. However, the following framework has been prepared:

1. **Event families defined by objective anchors** – e.g., (i) U.S. defence budget increases above a 5 % year-on-year threshold; (ii) CHIPS Act passage and semiconductor export restrictions; (iii) rare-earth export bans or trade sanctions; (iv) major cyber-attacks on critical infrastructure.
2. **Theme basket construction** – within each event date, the classifier groups securities by theme and forms equal-weighted baskets (e.g., all prime defence & aerospace stocks). Benchmarks such as SPY or sector ETFs are used for relative returns.
3. **Forward return calculations** – compute 126-day (6-month) and 252-day (12-month) returns after each event onset and compare against baseline periods. Use bootstrapped confidence intervals and placebo dates for robustness.
4. **Reporting** – summarise median and interquartile range of relative returns and highlight themes that historically benefitted or underperformed around such events.

Once real price data are connected, the module can be activated to produce tables such as “In the past four defence-budget expansions (1998–2025), the defence basket delivered a median +X % 6-month excess return vs SPY with an interquartile range of Y %.” Until then, the placeholder in `research_system.py` signals where to insert the analysis.

## Limitations and next steps

1. **Synthetic features** – all trend/volatility/liquidity measures are random numbers. Real historical price and volume data must be ingested (e.g., via a market data API or local database) to compute meaningful features, drawdowns and liquidity metrics.
2. **Limited classification recall** – because only keyword rules were used and no industry codes were available, many relevant companies (e.g., defence subcontractors or material processors with obscure names) remain unlabelled. Incorporating NAICS/GICS codes and natural-language embedding models would improve coverage and confidence scores.
3. **Fundamental data missing** – market capitalisation, shares outstanding and debt metrics were unavailable. These should feed into risk flags (e.g., dilution risk) and the scoring model.
4. **No real historical parallels** – the event analysis awaits price histories. Defining objective event triggers and performing a walk-forward backtest with purged cross-validation will be critical to avoid look-ahead bias.

## Implementation notes for automation

The provided `research_system.py` script is self-contained and can be scheduled in a Windows PowerShell workflow. Key operational points:

1. **Data refresh** – replace the environment variables `NASDAQ_LISTED_PATH` and `OTHER_LISTED_PATH` with the actual path to the daily listing feeds or remove them to download from the official NasdaqTrader site. The script writes timestamped CSVs (`universe_YYYYMMDD.csv`, `classified_YYYYMMDD.csv`, etc.) into the working directory.
2. **Price/volume ingestion** – integrate a module that reads daily OHLCV data from your preferred vendor. Ensure you honour API rate limits and cache results locally. The script currently assumes 252 trading days; adjust `MIN_HISTORY_DAYS` if necessary.
3. **Risk flags** – update `compute_risk_flags()` to include micro-cap hazards, zero-volume days and dilution proxies when fundamental data are available.
4. **Monitoring outputs** – the final CSVs (`scored_*.csv` and `eligible_*.csv`) are ready to load into Excel/Power BI. Each row includes the symbol, theme labels, all features, risk flags, eligibility and the monitor priority score. The `gate_fail_reasons` column enumerates why a security was excluded.

## What would change my mind?

This monitor system is designed to **evolve with data and policy developments**. Factors that would alter the screening priorities include:

1. **Major policy actions** – e.g., a new defence procurement bill, CHIPS Act II, or a critical-materials nationalisation programme. Such events would recalibrate the event module and theme weights.
2. **Technological breakthroughs** – if a disruptive technology (e.g., quantum computing, fusion energy) emerges within the 6–12 month horizon, a new theme should be added with its own classification rules and risk indicators.
3. **Data quality improvements** – access to comprehensive fundamentals or intraday liquidity metrics could lead to refined risk flags and gating thresholds.
4. **Empirical evidence** – once historical parallels are computed, empirical distributions may show that certain themes consistently underperform after specific events, prompting a downweighting in the scoring rubric.

Overall, this run demonstrates the **architecture and workflow** for a defensible monitoring system. With real data and richer classifications, it can become a powerful tool for tracking how defence, strategic technology and critical materials names respond to policy and geopolitical shifts.

---

1 GAO-24-107176, Critical Materials: Action Needed to Implement Requirements That Reduce Supply Chain Risks

<https://www.gao.gov/assets/gao-24-107176.pdf>

2 The CHIPS Program Office Vision for Success: Two Years Later

<https://www.nist.gov/document/chips-america-vision-success-two-year-report>

3 Global Compute and National Security | CNAS

<https://www.cnas.org/publications/reports/global-compute-and-national-security>

4 Strengthening the Reliability and Security of the United States Electric Grid – The White House

<https://www.whitehouse.gov/presidential-actions/2025/04/strengthening-the-reliability-and-security-of-the-united-states-electric-grid/>

5 6 Building Secure Supply Chains for America's Energy Future

<https://www.energy.gov/sites/default/files/2025-01/critical-minerals-materials-program-january2025.pdf>

7 Addressing the Threat to National Security from Imports of Copper – The White House

<https://www.whitehouse.gov/presidential-actions/2025/02/addressing-the-threat-to-nationalsecurity-from-imports-of-copper/>

8 raw.githubusercontent.com

<https://raw.githubusercontent.com/datasets/nasdaq-listings/main/data/nasdaq-listed-symbols.csv>

9 raw.githubusercontent.com

<https://raw.githubusercontent.com/datasets/nyse-other-listings/main/data/other-listed.csv>

10 A Time to Act: Policies to Strengthen the US Robotics Industry | ITIF

<https://itif.org/publications/2025/07/18/time-to-act-policies-to-strengthen-us-robotics-industry/>

11 DHS warns of escalating threats to US critical infrastructure in 2025 Homeland Threat Assessment - Industrial Cyber

<https://industrialcyber.co/threat-landscape/dhs-warns-of-escalating-threats-to-us-critical-infrastructure-in-2025-homeland-threat-assessment/>