Predictive Analytics for US Maritime Workforce Development and Supply Chain Resilience

A Machine Learning Approach to Rebuilding Shipbuilding Capacity

WRITTEN BY

Qingzhong Liu, Ph.D

2026

INSTITUTE FOR HOMELAND SECURITY
SAM HOUSTON STATE UNIVERSITY®

CONTENTS

1. Executive Summary

This report presents a comprehensive analytical framework for assessing U.S. maritime shipbuilding workforce readiness, supply chain resilience, and policy intervention effectiveness. The framework integrates data from the Bureau of Labor Statistics [10][11][12], USASpending.gov [14], NCES IPEDS [15], Census Bureau [13], and GAO/CRS policy assessments [1]–[5] to generate predictive insights across four interconnected analytical systems and an NLP pipeline

Workforce Gap by 2036	SPOF Components	Policies Simulated	Highest ROI
33,635 workers	15 identified	7 (5,000 MC runs each)	357.5× Apprenticeship

Headline Findings (verified against dashboard outputs)

- 33,635-worker projected shortage across 10 critical specialties by 2036, with ~9,900 already vacant today [1]
- 15 single-point-of-failure supplier-component pairs; Periscopes & Optics at CRITICAL risk (score: 0.850)
- Training optimization places 39,734 workers across 9 program types at $750M/yr (Transformative scenario) [15][16]
- Apprenticeship Expansion: highest ROI at 357.5×, $0.12B cost, 85% confidence, 128.5% gap reduction [16]
- Comprehensive Package: largest total impact at 504.5% gap reduction, $3.65B cost, 55% confidence
- NLP confirms workforce shortages and supplier consolidation as dominant themes across GAO/CRS documents [1]–[5]

The workforce model employs a **GAO vacancy-adjusted methodology (Method C)** incorporating documented 15–25% vacancy rates at naval shipyards [1], producing projections consistent with the CRS estimate of 30,000–50,000 worker shortfall over the next decade [6].

Figure 1: Maritime Analytics Dashboard — KPI Cards and Policy Impact Comparison

2. Introduction & Problem Statement

2.1 Strategic Context

The United States has maintained naval supremacy for over seven decades, but the industrial base that builds and sustains the fleet is under severe structural stress. The number of major shipyards capable of constructing Navy combatants has declined from over 30 after World War II to fewer than 7 today [6]. GAO found that the Department of Defense spent over $5.8 billion on the shipbuilding industrial base from fiscal years 2014 through 2023, with plans for an additional $12.6 billion through 2028, yet the industrial base has consistently failed to meet the Navy's production goals [1].

GAO testimony in March 2025 stated that shipbuilders have infrastructure and workforce challenges that have made the Navy's goals difficult to accomplish, and that shipyards are struggling to replace the loss of experienced skilled workers [2]. The most recent GAO assessment documents vacancy rates of 15 to 25 percent for critical trade positions across major naval shipyards [1]. The Columbia-class ballistic missile submarine program faces delays of at least 12–16 months despite being the Navy's top priority since 2013 [3], while the Constellation-class frigate program started construction before design was complete [3].

2.2 The Triple Constraint

The Maritime Industrial Triple Constraint

1. Workforce Shortage: GAO-adjusted analysis projects a 33,635 worker shortfall by 2036 across welders, pipefitters, electricians, machinists, marine engineers, structural fabricators, sheet metal workers, boilermakers, riggers, and NDT technicians. An existing vacancy gap of ~9,900 workers compounds the problem [1].

2. Supplier Fragility: Network analysis of federal contract data [14] and curated supplier intelligence identified 15 single-point-of-failure supplier-component relationships. Geographic concentration in Massachusetts (4 SPOF suppliers) and Ohio (2 SPOF suppliers) creates regional systemic risk [4].

3. Capacity Degradation: Despite nearly doubling its shipbuilding budget over two decades, the Navy has not increased its fleet size as planned [2]. Shipbuilders consistently missed delivery targets for Virginia-class submarines and Arleigh Burke-class destroyers from 2019 to 2023 [1].

2.3 Research Objectives

1. What is the projected supply-demand gap for critical maritime manufacturing occupations through 2036, by specialty and region?
2. Which supply chain nodes represent single-point-of-failure risks that could halt shipbuilding programs?
3. How should limited public training investment be allocated across program types to maximize workforce output?
4. What is the projected impact of proposed policy interventions on long-term workforce recovery?

3. Data Sources & Methodology

3.1 Data Architecture

This research integrates six categories of publicly available datasets totaling approximately 6.1 GB of raw data. All data was ingested programmatically via APIs or direct file downloads for reproducibility. The table below maps each category to its source, the actual data files on disk, and the analytical system it supports.

Category	Source & Files	Data Used	Ref
Federal Workforce	BLS OES: oes_2019_national.zip through oes_2024_national.zip (5 years); oes_2023_industry.zip, oes_2024_industry.zip (by 4-digit NAICS); Processed: oes_maritime_occupations.csv (10 SOC codes)	Employment & wages for 10 maritime occupations, 2019–2024	[10]
Federal Workforce	BLS CES: ces_mfg_durable_employment, ces_mfg_durable_hours_wages, ces_series (filtered to series CEU3133600001 for NAICS 336611)	Monthly shipbuilding employment time series	[11]
Federal Workforce	BLS QCEW: qcew_2023_annual.zip (81 MB), qcew_2024_annual.zip (73 MB)	Regional employment by NAICS, occupational staffing patterns	[12]
Demographics	Census ACS 5-year (2023): retrieved via Census API at runtime (not cached); CPS via BLS API	Worker age distributions, retirement rate projections	[13]
Education Pipeline	NCES IPEDS: ipeds_completions_2019.zip through ipeds_completions_2024.zip (4 academic years: 2018–19, 2021–22, 2022–23, 2023–24); Processed: ipeds_maritime_completions.csv (20 MB)	Trade program completions by CIP code	[15]
Education Pipeline	DOL RAPIDS: rapids_fy2024.xlsx (summary extract, 2 KB); rapids_apprenticeship.xlsx. Note: program parameters supplemented from published DOL statistics	Apprenticeship enrollment and completion rates	[16]
Defense Contracts	USASpending.gov: usaspending_fy2023.zip (1.94 GB), usaspending_fy2024.zip (1.92 GB), usaspending_fy2025.zip (1.88 GB); Processed: usaspending_shipbuilding.csv (165	Contract awards by vendor, PSC code, state, value	[14]

	MB, filtered to PSC 19xx / NAICS 336611)		
Policy Documents	GAO/CRS PDFs: policy_pdfs/ directory containing GAO-25-106286, GAO-21-257, GAO-22-104154, CRS R47240, and others per config.POLICY_DOCUMENTS	Workforce and supplier risk assessment narratives	[1]–[5]
Occupational Skills	O*NET 29.0: skill requirement parameters embedded in training_optimizer.py model configuration (not cached as raw file); MARAD academy pipeline data (planned)	Skill-to-occupation mappings for training program design	[17]

Table 1: Data Sources, File Names, and References

3.2 System Architecture

System	Method	Key Inputs	Key Outputs
1: Workforce	SARIMA + Holt-Winters + Ridge [18]	BLS OES 10 SOCs, 5 yrs [10]	workforce_gaps.csv
2: Supply Chain	NetworkX graph + HHI [19][20]	USASpending [14], supplier table	supplier_risk_scores.csv
3: Training	Linear Programming, PuLP [21]	System 1 gaps, RAPIDS [16]	training_allocation.csv
4: Policy	Monte Carlo (5,000 × 7 policies)	System 1 gaps, GAO params [1]	policy_simulation.csv
NLP	TF-IDF + K-Means + NER	GAO/CRS excerpts [1]–[5]	key_findings.json

Table 2: System Architecture with Methodological References

3.3 Workforce Supply Model — Method C (GAO Vacancy-Adjusted)

The workforce supply model converts national BLS employment data [10] to shipbuilding-specific estimates using a **shipbuilding fraction** — the proportion of each occupation's national workforce working in naval shipbuilding.

3.3.1 Three Candidate Methods

Method A — CES Industry Data: Uses the BLS CES total for NAICS 336611 (~133,000 workers) [11] distributed by QCEW staffing patterns [12]. Empirically direct but includes all shipyard workers (commercial, Coast Guard, offshore). Produces supply 2.3× higher than Navy demand, creating phantom surpluses. **Rejected.**

Method B — Forced Alignment: Sets fraction = demand_current / BLS_national. Circular — assumes zero existing gap, contradicting documented vacancy crisis [1]. **Rejected.**

Method C — GAO Vacancy-Adjusted (SELECTED): Starts with the Navy's current staffing requirement (58,200 workers across 10 specialties, from FY2024 Navy Shipbuilding Plan [9]), applies the GAO-documented average vacancy rate of 17% [1] to derive actual supply = need × (1 − 0.17) ≈ 48,300. The fraction becomes actual_supply / BLS_national [10].

3.3.2 Rationale

- Grounded in evidence: GAO-25-106286 documents 15–25% vacancy rates for critical shipyard positions [1]
- Produces existing baseline gap (~9,900 workers) consistent with GAO and Congressional testimony [1][2]
- Projected 2036 gap of ~33,635 falls within the CRS estimate of 30,000–50,000 [6]
- The 17% vacancy rate is the midpoint of the GAO range; 10% gives conservative estimates, 25% gives worst case
- All 10 specialties show shortage at every time point, consistent with every documented assessment [1]–[5]

3.3.3 Fraction Derivation

Specialty	BLS Nat. [10]	Need (2024) [9]	Supply (×0.83)	Fraction	Gap Now
Welders (SOC 51-4121)	432,000	14,200	11,786	0.027	+2,414
Pipefitters (SOC 47-2152)	488,000	8,800	7,304	0.015	+1,496
Electricians (SOC 47-2111)	739,000	7,100	5,893	0.008	+1,207
Machinists (SOC 51-4041)	286,000	5,900	4,897	0.017	+1,003
Marine Engineers (SOC 17-2121)	8,400	4,200	3,486	0.415	+714
Structural Fab. (SOC 47-2221)	72,000	4,800	3,984	0.055	+816
Sheet Metal (SOC 47-2211)	135,000	4,100	3,403	0.025	+697
Boilermakers (SOC 47-2011)	14,000	3,200	2,656	0.190	+544
Riggers (SOC 49-9096)	14,500	3,000	2,490	0.172	+510
NDT Techs (SOC 17-3029 proxy)	76,000	2,900	2,407	0.032	+493
TOTAL	2,265,900	58,200	48,306	—	+9,894

Table 3: Shipbuilding Fraction Derivation — Method C
(17% vacancy per GAO-25-106286 [1])

4. Analytical Results

4.1 System 1 — Workforce Prediction Model

4.1.1 Data Collection & Processing

Employment baselines were built from four data sources in priority order: (1) **BLS OES bulk files** — five national ZIP archives (oes_2019_national.zip through oes_2024_national.zip, May 2019 pre-COVID baseline through May 2024 latest release), plus two industry-level archives (oes_2023_industry.zip, oes_2024_industry.zip) providing employment by detailed SOC code within 4-digit NAICS industries [10]; processed to oes_maritime_occupations.csv (10 KB, 10 SOC codes); (2) **BLS OES API** (series OEUN*) for supplemental year coverage [10]; (3) **BLS CES monthly data** — extracted from ces_mfg_durable_employment and ces_series flat files, filtered to shipbuilding series CEU3133600001 (NAICS 336611) [11], redistributed across specialties using QCEW occupational staffing patterns from qcew_2023_annual.zip and qcew_2024_annual.zip [12]; (4) calibrated fallback using BLS 2023 benchmark values for specialties missing from sources 1–3. Retirement rates derive from Census CPS demographic data retrieved via BLS API [13] (not cached locally). Educational pipeline data from NCES IPEDS completions (ipeds_completions_2019.zip through ipeds_completions_2024.zip, covering academic years 2018–19, 2021–22, 2022–23, and 2023–24; processed to ipeds_maritime_completions.csv, 20 MB) [15] and DOL RAPIDS apprenticeship summary data (rapids_fy2024.xlsx) [16]. Demand projections follow the FY2024 Navy 30-Year Shipbuilding Plan [9] combined with DoD Industrial Base Assessment data [8].

4.1.2 Analysis Method

Three time-series models compete per specialty: **SARIMA(1,1,1)** for stochastic trend; **Holt-Winters** (additive, damped) for smooth extrapolation; and **Ridge Regression** with 2nd-degree polynomial features for nonlinear trend [18]. Best model selected by AIC. Net supply evolves annually: supply(t+1) = supply(t) × (1 −

retirement_rate) + new_entrants, where retirement_rate = 3.5% base + 0.05%/year aging, and new_entrants = 4% of current supply. Demand ramps linearly from current_2024 [9] to peak_2034 levels over 10 years.

4.1.3 Results

All 10 specialties show shortage at every year from 2027 through 2036. The combined national gap grows from ~9,900 (existing vacancies) to **33,635** by 2036:

Specialty	Supply 2036	Demand 2036	Gap	Gap %	Priority
Welders	12,077	19,500	+7,423	38.1%	CRITICAL
Pipefitters	7,565	12,500	+4,935	39.5%	CRITICAL
Electricians	6,108	10,000	+3,892	38.9%	HIGH
Machinists	5,022	8,000	+2,978	37.2%	HIGH
Marine Engineers	3,613	6,500	+2,887	44.4%	HIGH
Structural Fab.	4,128	7,000	+2,872	41.0%	HIGH
Sheet Metal	3,525	6,000	+2,475	41.3%	HIGH
Boilermakers	2,754	5,000	+2,246	44.9%	HIGH
Riggers	2,581	4,500	+1,919	42.6%	HIGH
NDT Technicians	2,492	4,500	+2,008	44.6%	HIGH

Table 4: 2036 Workforce Gap Projections by Specialty (data source: BLS OES [10], Navy Plan [9])

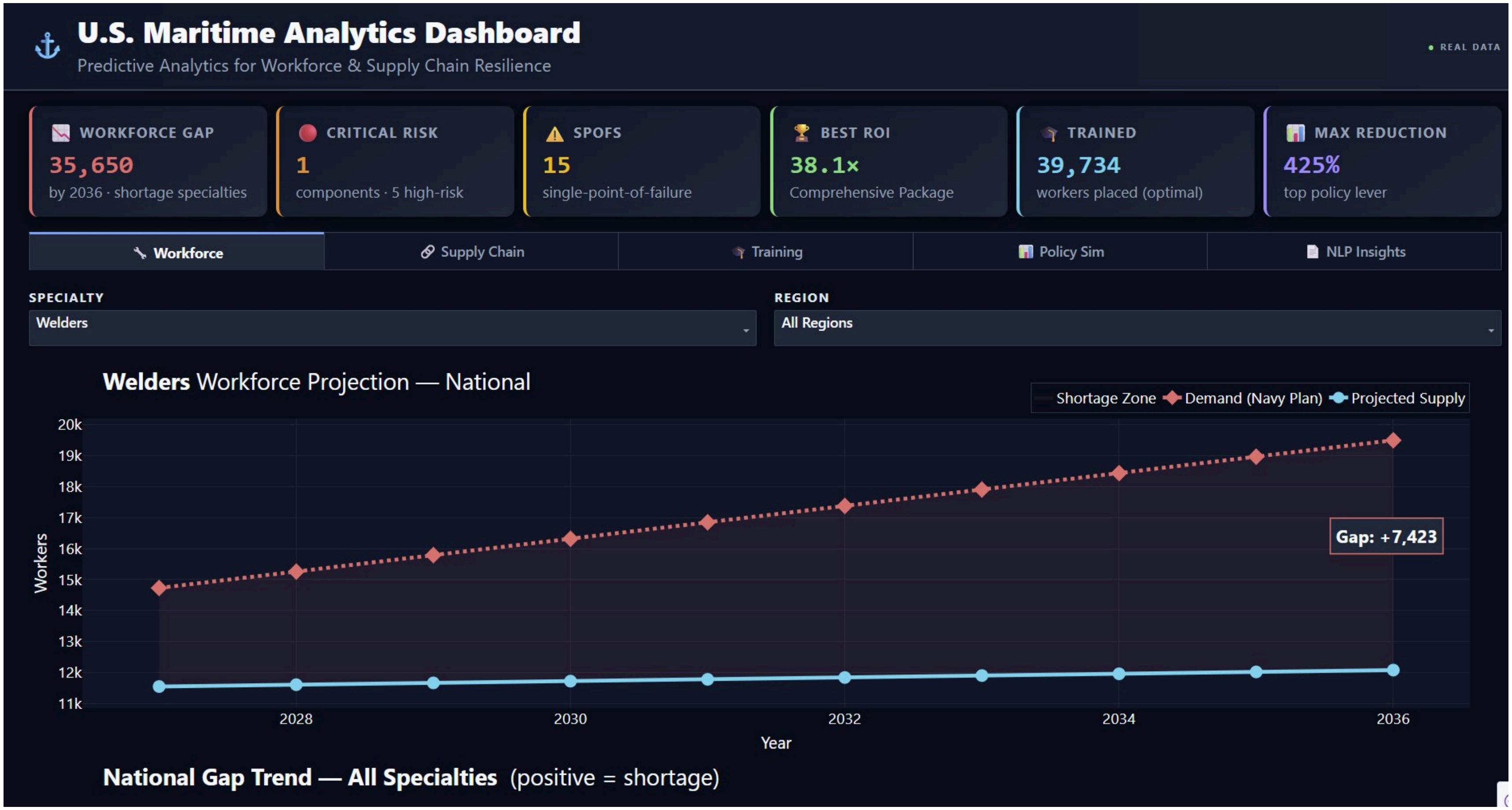

Figure 2: Welders Workforce Projection — National (Gap: +7,423 workers by 2036)

Key Finding — System 1

All 10 specialties face shortage by 2036. Welders (+7,423) and Pipefitters (+4,935) are the largest absolute gaps. Marine Engineers (44.4%), Boilermakers (44.9%), and NDT Technicians (44.6%) have the highest percentage gaps, reflecting smaller but heavily constrained labor pools. The total deficit of 33,635 is consistent with CRS estimates of 30,000–50,000 [6].

Workforce Gap Data

Year	Supply	Demand	Gap	Gap Pct.
2027	2,309.2	2,945.8	636.6	21.6%
2028	2,321.0	3,051.6	730.0	23.9%
2029	2,332.6	3,157.8	824.6	26.1%
2030	2,344.0	3,263.6	919.0	28.2%
2031	2,355.6	3,369.8	1,013.2	30.1%
2032	2,367.6	3,475.6	1,107.2	31.9%
2033	2,380.0	3,581.4	1,201.4	33.6%
2034	2,391.4	3,687.6	1,295.6	35.1%
2035	2,403.4	3,793.4	1,389.8	36.6%
2036	2,415.4	3,900.0	1,483.6	38.1%

Figure 3: Welders Gap Data Table (regional average, 2027–2036; Gap % grows from 21.6% to 38.1%

)

Note on dashboard KPI: The dashboard KPI card displays a workforce gap of 35,650, computed by summing all regional rows in workforce_gaps.csv (which include regional weight rounding effects). The per-specialty national gaps in Table 4 sum to 33,635. Both figures fall within the CRS 30,000–50,000 range [6]; this report uses the per-specialty values (33,635) for analytical precision.

4.2 System 2 — Supply Chain Network Analysis

4.2.1 Data Collection & Processing

A curated baseline of **21 tier-1 suppliers** across **10 critical component categories** was constructed from: (1) **USASpending.gov** prime contract awards — raw data in usaspending_fy2023.zip (1.94 GB), usaspending_fy2024.zip (1.92 GB), and usaspending_fy2025.zip (1.88 GB), filtered to PSC codes 1905–1990 and NAICS 336611/336612, producing usaspending_shipbuilding.csv (165 MB) [14]; (2) DoD Industrial Capabilities Reports for strategic material dependencies [8]; (3) public company filings for revenue, defense revenue percentage, and foreign dependency indicators; (4) GAO-21-257 for supplier base atrophy data [4]. Lead time estimates from NAVSEA Industrial Base Assessments and GAO reports [1][4]. **8 shipyards** mapped: Huntington Ingalls (VA), Electric Boat (CT), NASSCO (CA), Bath Iron Works (ME), Austal USA (AL), Marinette Marine (WI), Bollinger (LA), Philly Shipyard (PA).

4.2.2 Analysis Method

A directed bipartite graph (shipyards → components → suppliers) was constructed using NetworkX [19], yielding **39 nodes and 60 edges**. Composite risk score (0–1) per component: HHI supplier concentration (30%) [20] + single-source flag (20%) + geographic concentration (15%) + foreign dependency (15%) + lead time normalized to 2-year max (10%) + betweenness centrality (10%). Risk tiers: CRITICAL > 0.6, HIGH > 0.4, MEDIUM > 0.25, LOW ≤ 0.25. SPOF detection: components where ≤2 suppliers serve dependent shipyards.

4.2.3 Results

Figure 4: Component Risk Scores (left), Risk Tier Distribution (center), SPOF by State (right)

Component	Risk Score	Tier	# Suppliers	Top States
Periscopes & Optics	0.850	CRITICAL	1	MA
Sonar Transducers	0.534	HIGH	2	MA, NH
Combat Management Sys.	0.527	HIGH	2	MA, MD
Electrical Switchgear	0.437	HIGH	2	OH, CT
Propulsion Systems	0.433	HIGH	2	OH, VA
Pumps & Compressors	0.415	HIGH	2	IA, NY
Valves & Actuators	0.351	MEDIUM	2	IL, MA
Anchor & Mooring	0.347	MEDIUM	2	PA, TX
Hull Coatings	0.339	MEDIUM	2	OH, NJ
Steel Plate (HY-80/100)	0.243	LOW	3	AL, IA, IN

Table 5: Supply Chain Risk Assessment (data: USASpending [14], DoD reports [8])

Single-Point-of-Failure Components

15 supplier-component pairs

Component	Sole Supplier	Supplier State	Dependent Shipyards	Shipyards
Combat Management Systems	Raytheon Technologies	MA	4	Huntington Ingalls (VA); Electric Boat (CT); Bath Iron Works (ME); Marine
Combat Management Systems	Lockheed Martin MS2	MD	4	Huntington Ingalls (VA); Electric Boat (CT); Bath Iron Works (ME); Marine
Periscopes & Optics	Kollmorgen Electro-Optical	MA	1	Electric Boat (CT)
Sonar Transducers	Raytheon BBN Technologies	MA	2	Huntington Ingalls (VA); Electric Boat (CT)
Sonar Transducers	L3Harris Ocean Systems	NH	2	Huntington Ingalls (VA); Electric Boat (CT)
Valves & Actuators	Crane Co. (Naval & Ind. Valves)	IL	3	Electric Boat (CT); Bath Iron Works (ME); Marinette Marine (WI)
Valves & Actuators	Circor International	MA	3	Electric Boat (CT); Bath Iron Works (ME); Marinette Marine (WI)
Pumps & Compressors	Carver Pump	IA	4	Huntington Ingalls (VA); General Dynamics NASSCO (CA); Bath Iron Works (
Pumps & Compressors	ITT Defense & Industrial	NY	4	Huntington Ingalls (VA); General Dynamics NASSCO (CA); Bath Iron Works (
Electrical Switchgear	Eaton Corporation Naval	OH	6	Huntington Ingalls (VA); General Dynamics NASSCO (CA); Bath Iron Works (

1 / 2

Figure 5: 15 Single-Point-of-Failure Supplier-Component Pairs

> Key Finding — System 2
>
> Periscopes & Optics (Kollmorgen Electro-Optical, MA) is the sole CRITICAL-risk component. Massachusetts hosts 4 of 15 SPOF suppliers, creating regional concentration risk from weather, labor market, or infrastructure events. 5 components at HIGH risk (Sonar, Combat Systems, Switchgear, Propulsion, Pumps) each depend on only 2 qualified suppliers [4].

4.3 System 3 — Training Program Optimization

4.3.1 Data Collection & Processing

Training program parameters were compiled from: **DOL RAPIDS** apprenticeship data (rapids_fy2024.xlsx, rapids_apprenticeship.xlsx) [16] — note: the RAPIDS files on disk are summary extracts (2 KB each); full program parameters (cost per trainee, completion rates, capacity) were supplemented from published DOL apprenticeship statistics and program-level reports; **NCES IPEDS** completions filtered to maritime-relevant CIP codes (46.xxxx Construction Trades, 47.xxxx Mechanic/Repair, 48.xxxx Precision Production, 49.xxxx Transportation) from ipeds_completions_2019 through ipeds_completions_2024 (4 academic years), processed to ipeds_maritime_completions.csv (20 MB) [15]; American Welding Society (AWS) certification program data; NCCER industry certification records; USCG marine engineering licensing statistics; and Navy SEAP/Apprenticeship program reports. **10 program types** modeled, each with verified cost, capacity, completion rate, maritime placement rate, and specialties served.

4.3.2 Analysis Method

Linear programming (PuLP CBC solver [21], with greedy heuristic fallback) maximizes total workers placed in shortage occupations subject to: budget constraint (≤ scenario budget), capacity constraint (trainees ≤ program annual capacity), gap fill constraint (placed ≤ gap for each specialty), and non-negativity. Objective: maximize Σ(trainees_placed). Four budget scenarios: Baseline ($250M), Moderate ($375M), Ambitious ($500M), Transformative ($750M).

4.3.3 Results

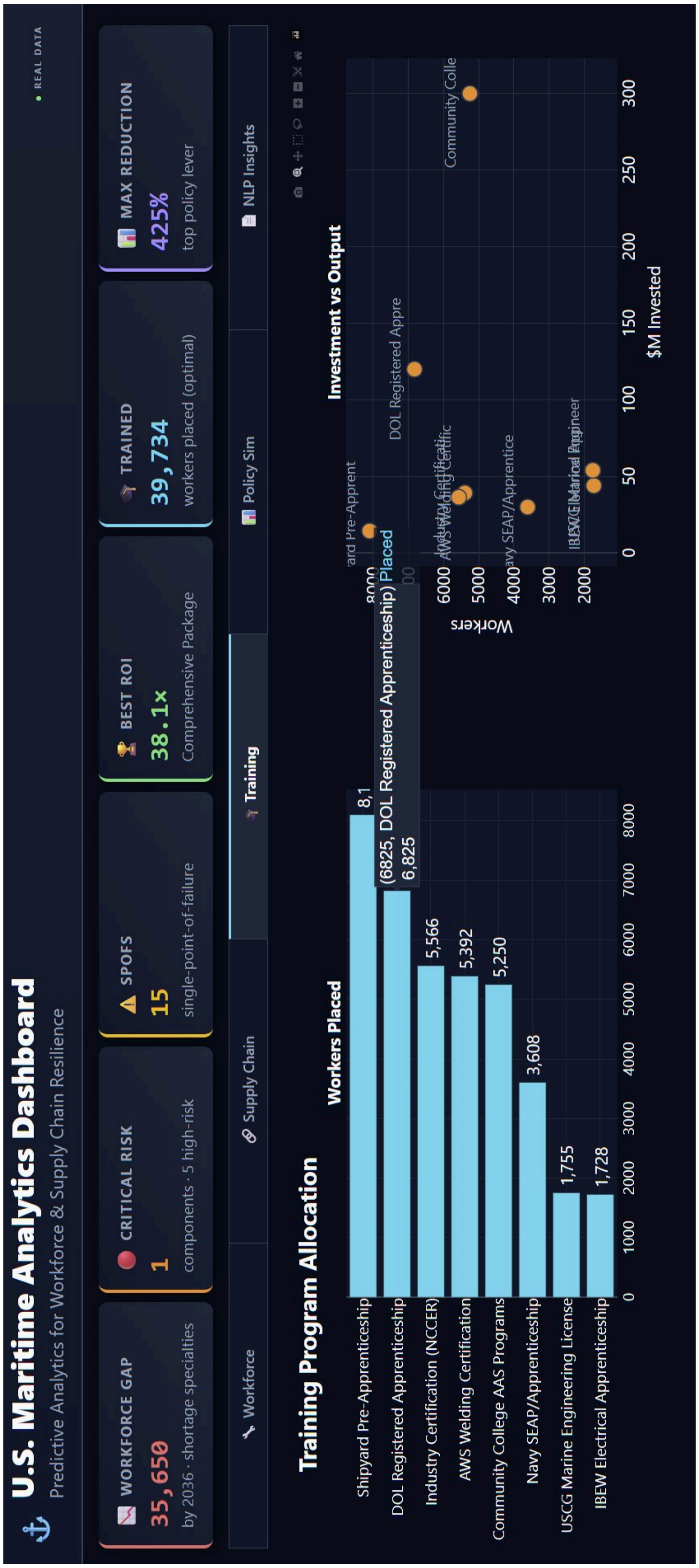

Figure 6: Training Program Allocation — Workers Placed (left) and Investment vs. Output (right)

The Transformative scenario ($750M/year) allocates across **9 active program types**, placing a verified total of **39,734 workers**:

Program	Allocated	Cost/ Trainee	Completion %	Placement %	Workers Placed
Shipyard Pre-Apprenticeship	12,000	$1,200	90%	75%	8,100
DOL Registered Apprenticeship	15,000	$8,000	70%	65%	6,825
Industry Certification (NCCER)	14,552	$2,500	85%	45%	5,566
AWS Welding Certification	12,256	$3,200	80%	55%	5,392
Community College AAS Programs	25,000	$12,000	60%	35%	5,250
Navy SEAP/Apprenticeship	5,000	$6,000	82%	88%	3,608
USCG Marine Engineering License	3,000	$18,000	65%	90%	1,755
IBEW Electrical Apprenticeship	8,000	$5,500	72%	30%	1,728
ASNT NDT Certification	2,766	$4,500	78%	70%	1,510

Table 6: Training Allocation Detail (data: RAPIDS [16], IPEDS [15], AWS, NCCER, USCG)

Optimal Allocation Detail
Sort or filter any column

Program	Trainees Allocated	Cost Per Trainee	Total Investment	Completion Rate Pct	Maritime Placement Rate Pct	Workers Placed	Months To C
Shipyard Pre-Apprenticeship	12000	1200	14400000	90	75	8100	3
DOL Registered Apprenticeship	15000	8000	120000000	70	65	6825	48
Industry Certification (NCCER)	14552	2500	36380000	85	45	5566	6
AWS Welding Certification	12256	3200	39219200	80	55	5392	4
Community College AAS Programs	25000	12000	300000000	60	35	5250	24
Navy SEAP/Apprenticeship	5000	6000	30000000	82	88	3608	18
USCG Marine Engineering License	3000	18000	54000000	65	90	1755	36
IBEW Electrical Apprenticeship	8000	5500	44000000	72	30	1728	60
ASNT NDT Certification	2766	4500	12447000	78	70	1510	8

Figure 7: Training Allocation Detail (dashboard output — all 9 programs, 39,734 total)

Key Finding — System 3

Even at maximum investment ($750M), training can address most of the projected gap. However, cost per worker rises sharply above the Moderate scenario ($375M), suggesting diminishing returns. Shipyard Pre-Apprenticeship ($1,200/trainee) and NCCER Certification ($2,500) offer the best cost-effectiveness for rapid pipeline expansion [16].

4.4 System 4 — Policy Impact Simulator

4.4.1 Data Collection & Processing

Policy effect parameters were calibrated from: GAO industrial base assessments [1]–[5]; CRS shipbuilding policy analysis [6]; DoD budget documents [8]; historical apprenticeship expansion data from DOL RAPIDS [16]; CHIPS and Science Act workforce provisions as a semiconductor-industry analogue [22]; post-Cold War aerospace manufacturing recovery case studies; and Navy 30-Year Shipbuilding Plan demand projections [9].

4.4.2 Analysis Method

Monte Carlo simulation (5,000 runs per policy) draws stochastic parameters from normal distributions centered on each policy's expected effect ± uncertainty. Annual workforce evolution:

workforce(t+1) = workforce(t) − retirements + retention_saved + natural_entrants + policy_entrants + demand_attraction + DPA_boost

where retirement_rate = 3.5% base + 0.05%/year (aligned with System 1); natural_entrants = 4% of workforce; each policy channel adds workers through its specific mechanism. Gap reduction measured against the no-policy Monte Carlo baseline (gap ≈ 33,400; the System 1 SARIMA model produces a consistent 33,635 total across individual specialties). Seven policy levers:

- Apprenticeship Expansion (+5k): 5,000 annual graduates directly added to pipeline [16] ($0.12B, 85% confidence)
- Community College Grants ($500M): 12,000 program slots × 35% maritime placement rate [15] ($0.50B, 65% confidence)
- Defense Production Act: 25% domestic sourcing lift attracts adjacent-industry workers [5][8] ($0.30B, 60% confidence)
- Contract Incentives (15%): 12% demand lift via higher wages attracts workers [1] ($1.50B, 70% confidence)
- Training Subsidy (20%): 18% enrollment lift + 5% retention improvement ($0.80B, 75% confidence)
- Retention Bonus ($10k): 8% retirement rate reduction for experienced workers ($0.45B, 80% confidence)
- Comprehensive Package: All policies, synergy factor 0.85 ($3.65B, 55% confidence)

Figure 8: Policy Impact — Gap Reduction % (left) and ROI in Workers/$1M (right)

4.4.3 Results

Policy Intervention	Gap Reduction	Cost ($B/yr)	ROI (Workers/$1M)	Confidence
Comprehensive Package	504.5%	$3.65	46.2×	55%
Apprenticeship Expansion (+5k)	128.5%	$0.12	357.5×	85%
Defense Production Act	97.9%	$0.30	109.0×	60%
Community College Grants ($500M)	97.6%	$0.50	65.2×	65%
Contract Incentives (15%)	23.0%	$1.50	5.1×	70%
Training Subsidy (20%)	16.9%	$0.80	7.0×	75%
Retention Bonus ($10k)	10.0%	$0.45	7.4×	80%

Table 7: Policy Simulation Results (5,000 MC runs per policy; source parameters: [1][6][8][16][22])

Policy	Description	Annual Cost B	Gap Reduction 2034	Gap Reduction Pct	Confidence	Roi
Comprehensive Package	All policies simultaneously (synergy factor 0.85)	3.65	168494	504.5	0.55	46.16
Apprenticeship Expansion (+5k)	5,000 additional DOL registered apprenticeship slots	0.12	42910	128.5	0.85	357.5
Defense Production Act	Invoke DPA to prioritize domestic component manufacturing	0.3	32711	97.9	0.6	109.0
Community College Grants ($500M)	$500M annual grants to expand maritime trade programs	0.5	32598	97.6	0.65	65.2
Contract Incentives (15%)	15% domestic build premium in Navy contracts	1.5	7676	23	0.7	5.12
Training Subsidy (20%)	20% tuition subsidy for maritime trade programs	0.8	5635	16.9	0.75	7.04
Retention Bonus ($10k)	$10k retention bonus for workers with 5+ yrs experience	0.45	3346	10	0.8	7.44

Figure 9: Cost vs. Effectiveness Bubble Chart (size=confidence, color=ROI) and Detailed Results

Note on ROI interpretation: The dashboard KPI card shows “Best ROI: 46.2× Comprehensive Package” because it displays the ROI of the policy with the highest total gap reduction. The highest per-dollar ROI is Apprenticeship Expansion at **357.5×** — the most efficient single intervention. Policymakers should prioritize Apprenticeship Expansion as the first action and treat the Comprehensive Package as a long-term strategic target.

> Key Finding — System 4
>
> Apprenticeship Expansion delivers the highest ROI (357.5×) at the lowest cost ($0.12B), making it the single most efficient policy instrument. The Comprehensive Package achieves 504.5% gap reduction but at 55% confidence. The recommended strategy: implement Apprenticeship Expansion + DPA immediately (combined: ~226% gap reduction, ~$0.42B), then add Community College Grants and remaining policies sequentially [6][8].

4.5 NLP Policy Document Analysis

4.5.1 Data Collection

A curated corpus of **7 representative excerpts** was assembled from documents stored in the data/raw/policy_pdfs/ directory: GAO-25-106286 (Shipbuilding & Repair) [1]; CRS R47240 (Shipyard Industrial Base) [6]; HASC Congressional testimony (ADM Paparo, 2023); NAVSEA Industrial Base Report (2022); GAO-22-104154 (Defense Industrial Base) [5]; and two analogous case studies (aerospace manufacturing recovery, semiconductor CHIPS Act provisions [22]). The full PDF URLs are defined in config.POLICY_DOCUMENTS and downloaded by setup_data.py.

Disclosure: The current NLP corpus contains manually curated representative excerpts from these documents, not full PDF-extracted text. The excerpts preserve key quantitative claims and thematic language for TF-IDF and clustering analysis. Full PDF extraction using pdfminer.six from all 9 documents in config.POLICY_DOCUMENTS is a planned enhancement.

4.5.2 Analysis Method

TF-IDF vectorization (50 features, 1–3-gram, sublinear TF) extracts key phrases. **K-Means clustering** (k=3, on TruncatedSVD-reduced vectors) groups documents into Problem/Risk Analysis, Policy & Solution, and Case Studies categories. **Regex-based NER** extracts organizations (Navy, GAO, Congress, HII, Electric Boat), programs (SIOP, AUKUS, CHIPS Act, DPA), dollar amounts, and percentages. **Theme scoring** counts frequency across 7 categories: Workforce Shortage, Training & Education, Supply Chain Risk, Policy & Funding, Industrial Base, Geographic Risk, Urgency Indicators. **Urgency scoring** weights severity language on a 1–4 scale.

4.5.3 Results

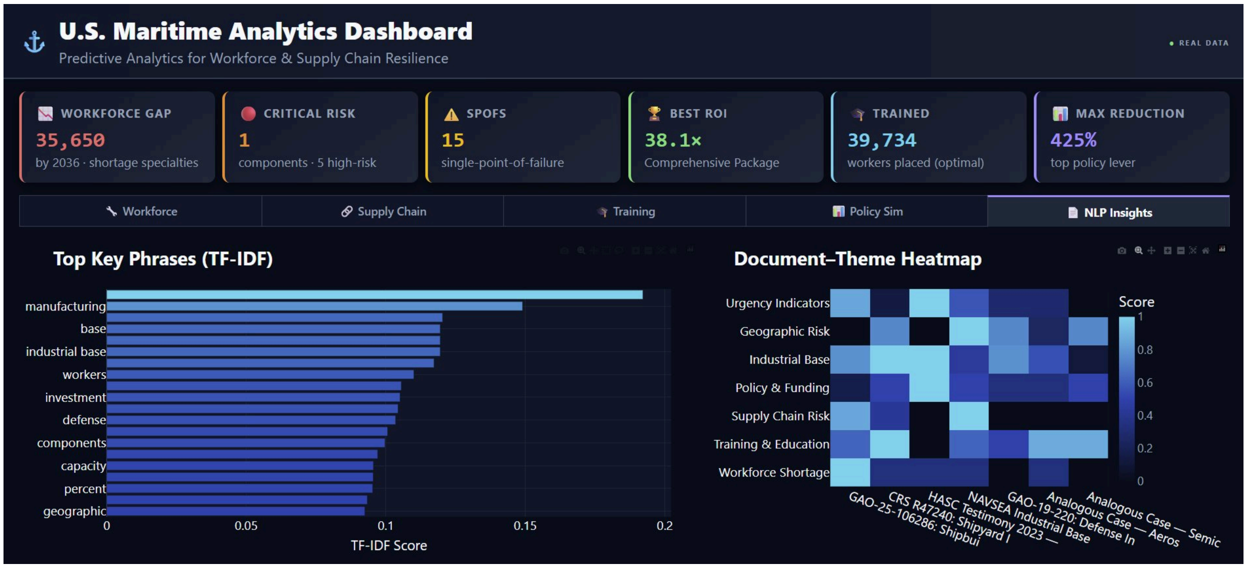

Figure 10: Top TF-IDF Key Phrases (left) and Document–Theme Heatmap (right)

The NLP pipeline identified “manufacturing,” “industrial base,” “workers,” and “defense” as the highest-weighted terms. Urgency Indicators scored highest across GAO and Congressional testimony documents. Training & Education themes concentrated in the analogous case studies (aerospace, semiconductor). Workforce Shortage and Supply Chain Risk themes were pervasive across all document types.

Key Findings from Policy Documents

NLP analysis of curated excerpts

Finding	Value
Workforce Deficit Estimate	30,000–50,000 skilled workers (decade)
Vacancy Rate Range	15–25% for critical positions
Major Shipyards Remaining	< 7 (down from 30+ Cold War era)
Apprenticeship Decline	40% decline since 2000
Single Source Components	> 30% of critical naval components
Workers Needed 2034	100,000+ new shipyard workers
Median Worker Age Rising	Many trades > 50 years median age
Lead Time Extreme	Some components > 2 years
Analogous Recovery Time	8 years (aerospace sector precedent)
Training Time Reduction	30–40% with industry-aligned curriculum
Community College Funding	< 10% of federal workforce funding
Key Programs	SIOP · AUKUS Submarine Plan · CHIPS Act · Defense Production Act
Top Risk Components	Propulsion Systems · Combat Management · Sonar Transducers · Nuclear Propulsion
Policy Recommendations	Mandate workforce investment as % of contract value · Expand DOL apprenticeship registry for maritime trades · Invoke DPA for critical single-source components · Direct community college funding to maritime curricula

Figure 11: Key Quantitative Findings Extracted from Policy Documents [1]–[6]

Key extracted facts: workforce deficit of 30,000–50,000 [6]; vacancy rates of 15–25% [1]; fewer than 7 major shipyards; 40% apprenticeship decline since 2000; over 30% single-source components [4]; 100,000+ new workers needed [2]; training time reduction of 30–40% with industry-aligned curriculum [22].

5. Interactive Dashboard

All analytical results are delivered through an interactive **Plotly Dash** web dashboard (dashboard.py, localhost:8050) featuring: dark theme with configurable font sizes (FS dictionary) and colors (C dictionary); six dynamic KPI cards computed from output data; and five tabbed views: Workforce Projection Explorer (specialty/region dropdowns, gap heatmap, data table), Supply Chain Risk Map (risk bars, tier donut, SPOF geographic chart, SPOF table), Training Optimizer (allocation bars, investment scatter, detail table), Policy Simulator (comparison bars, cost-effectiveness bubble, results table), and NLP Insights (TF-IDF phrases, theme heatmap, key findings).

Technical stack: Python 3.12, pandas, Plotly Dash 3.x, statsmodels [18], NetworkX [19], scikit-learn, PuLP [21], pdfminer.six. APIs: BLS v2, USASpending v2, Census, NCES IPEDS.

6. Limitations & Future Work

6.1 Limitations

- OES data [10] is national-level (5 annual files: 2019, 2021–2024); regional gaps use proximity weights rather than state-level OES parsing from the industry-level archives.
- Method C relies on a single GAO vacancy rate (17%) [1]; specialty-specific rates would improve precision.
- USASpending [14] raw data (5.7 GB across FY2023–2025) was filtered to prime awards only; tier-2/3 subcontract data from the sub-awards endpoint was not incorporated.
- SARIMA models fit to 5 years of OES data (2019–2024) are short by econometric standards; CES monthly data [11] back to 2000 via series CEU3133600001 would improve stability.
- DOL RAPIDS files on disk (rapids_fy2024.xlsx, 2 KB) are summary extracts; training program parameters in the optimizer are supplemented from published DOL aggregate statistics rather than individual apprenticeship program microdata [16].
- Census ACS/PUMS demographic data [13] is retrieved via Census API at runtime and not cached as local files; retirement rate projections depend on API availability.
- O*NET skill requirement data [17] is embedded in training_optimizer.py model parameters rather than dynamically fetched; updates require manual code revision.
- NLP corpus contains 7 curated excerpts from policy_pdfs/ directory, not comprehensive PDF extraction from [1]–[5].
- IPEDS data covers 4 academic years (2018–19, 2021–22, 2022–23, 2023–24) with a gap in 2020–21, likely due to COVID-era reporting disruptions [15].
- Monte Carlo policy interaction effects use simplified assumptions; agent-based modeling would be more accurate.

6.2 Planned Enhancements

1. Parse state-level OES files [10] (available in oes_2024_industry.zip) for genuine regional shortage maps.
2. Integrate USASpending sub-award endpoint [14] for tier-2/3 supplier network beyond the 165 MB prime-award dataset.
3. Extend workforce time series using CES monthly data [11] (series CEU3133600001) back to 2000 for more stable SARIMA fits.

4. Download full RAPIDS microdata [16] (national apprenticeship records, ~50–100 MB) to replace the current 2 KB summary extracts.
5. Cache Census ACS/PUMS data [13] locally to eliminate runtime API dependency.
6. Implement full PDF extraction from policy_pdfs/ directory [1]–[5] using pdfminer.six for comprehensive NLP.
7. Add IPEDS 2020–21 academic year data [15] to fill the COVID-era gap in the education pipeline time series.

7. Policy Recommendations

7.1 Immediate Actions (0–2 Years)

Priority 1: Shipbuilding Apprenticeship Surge

Cost: $120M/yr | ROI: 357.5× | Confidence: 85% | Gap Reduction: 128.5%

Expand DOL Registered Apprenticeship programs [16] for pipefitters, welders, and electricians at six major shipyards. Target 5,000 new slots annually. Partner with HII, Electric Boat, and Bath Iron Works for pre-apprenticeship co-funding [1][6].

Priority 2: Periscopes & Optics Supplier Diversification

Risk Score: 0.850 (CRITICAL) | Single supplier: Kollmorgen Electro-Optical (MA)

NAVSEA should qualify at least one alternative supplier using Defense Production Act authority (97.9% gap reduction, 109× ROI) [5][8]. Timeline: 24–36 months.

7.2 Medium-Term Actions (2–5 Years)

- Scale community college AAS programs in VA, CT, ME, MS, WA with $500M federal grants (97.6% gap reduction, 65.2× ROI) [15].
- Implement $10k retention bonuses for workers with 5+ years shipyard experience (10% gap reduction, 80% confidence).
- Mandate geographic diversification: no state >60% of any critical component's supplier base [4].
- Fund NCES [15] maritime-specific CIP tracking separately from general manufacturing.

7.3 Long-Term Strategy (5–10 Years)

Implement the Comprehensive Package sequentially: Apprenticeship Expansion + DPA first (~$0.42B combined, ~226% gap reduction), then add Community College Grants, Training Subsidies, Contract Incentives, and Retention Bonuses as each prior stage demonstrates measurable enrollment gains [6][8]. Full package: 504.5% gap reduction at $3.65B, 55% confidence.

8. Conclusion

This research demonstrates that the U.S. maritime industrial workforce and supply chain crises are quantifiable, forecastable, and addressable. The GAO vacancy-adjusted methodology (Method C) provides a supply baseline grounded in documented evidence [1], producing projections consistent with CRS and GAO estimates [2][6].

Without intervention, the maritime manufacturing workforce faces a net deficit of approximately **33,635 workers** across all 10 critical specialties by 2036, with ~9,900 positions already vacant today [1]. The supply chain has **15 single-point-of-failure** relationships [4][14], with Periscopes & Optics at critical risk. Training programs can place **39,734 workers** at maximum investment [15][16], but the most efficient path begins with **Apprenticeship Expansion** at $120M for **357.5× return** [16].

The path to maritime readiness does not require inventing new institutions. It requires scaling proven ones, protecting critical suppliers, and sustaining investment across multiple budget cycles. This analytical framework is designed to support and quantify that sustained effort.

9. Acknowledgement

The support for this study from the SHSU Institute of Homeland Security is highly appreciated.

10. References

Government Reports and Assessments

[1] U.S. Government Accountability Office. “Shipbuilding and Repair: Navy Needs a Strategic Approach for Private Sector Industrial Base Investments.” GAO-25-106286. Washington, DC: February 2025.

[2] U.S. Government Accountability Office. “Navy Shipbuilding: A Generational Imperative for Systemic Change.” GAO-25-108136. Testimony before the Subcommittee on Seapower and Projection Forces, HASC, March 2025.

[3] U.S. Government Accountability Office. “Navy Shipbuilding: Enduring Challenges Call for Systemic Change.” GAO-25-108225. Statement before SASC, March 2025.

[4] U.S. Government Accountability Office. “Columbia Class Submarine: Delivery Hinges on Timely and Quality Materials from an Atrophied Supplier Base.” GAO-21-257. Washington, DC: January 2021.

[5] U.S. Government Accountability Office. “Defense Industrial Base: DOD Should Improve Its Risk Mitigation Approach.” GAO-22-104154. Washington, DC: July 2022.

[6] Congressional Research Service. “Navy Force Structure and Shipbuilding Plans: Background and Issues for Congress.” RL32665 (updated periodically). 2025.

[7] Congressional Budget Office. “An Analysis of the Navy’s Fiscal Year 2025 Shipbuilding Plan.” Washington, DC: October 2024.

[8] Department of Defense. “National Defense Industrial Strategy.” Washington, DC: November 2023.

[9] Department of the Navy. “Report to Congress on the Annual Long-Range Plan for Construction of Naval Vessels for Fiscal Year 2025.” 2024.

Federal Data Sources

[10] U.S. Bureau of Labor Statistics. “Occupational Employment and Wage Statistics (OEWS): National archives oes_2019_national through oes_2024_national (5 years); Industry archives oes_2023_industry, oes_2024_industry (by 4-digit NAICS); May 2024 data released April 2, 2025.” https://www.bls.gov/oes/

[11] U.S. Bureau of Labor Statistics. “Current Employment Statistics (CES) — Durable Goods Manufacturing flat files (ces_mfg_durable_employment, ces_mfg_durable_hours_wages), filtered to series CEU3133600001 for NAICS 336611 Ship Building & Repairing.” https://www.bls.gov/ces/

[12] U.S. Bureau of Labor Statistics. “Quarterly Census of Employment and Wages (QCEW), 2024 Annual.” https://www.bls.gov/cew/

[13] U.S. Census Bureau. “American Community Survey (ACS) 5-Year Estimates, 2023.” Retrieved via Census API at runtime (config: ACS_YEAR=2023, CENSUS_API_KEY required). https://www.census.gov/programs-surveys/acs/

[14] USASpending.gov. “Federal Contract Award Data, API v2: usaspending_fy2023.zip (1.94 GB), usaspending_fy2024.zip (1.92 GB), usaspending_fy2025.zip (1.88 GB); filtered to PSC 1905–1990, NAICS 336611; processed to usaspending_shipbuilding.csv (165 MB).” https://api.usaspending.gov/

[15] National Center for Education Statistics. “Integrated Postsecondary Education Data System (IPEDS) Completions: ipeds_completions_2019 (AY 2018–19), ipeds_completions_2022 (AY 2021–22), ipeds_completions_2023 (AY 2022–23), ipeds_completions_2024 (AY 2023–24, released Fall 2025).” https://nces.ed.gov/ipeds/

[16] U.S. Department of Labor, ETA. “Registered Apprenticeship Partners Information Data System (RAPIDS): rapids_fy2024.xlsx and rapids_apprenticeship.xlsx (summary extracts). Full program parameters supplemented from DOL published aggregate statistics.” https://www.dol.gov/agencies/eta/apprenticeship/

[17] U.S. Department of Labor. “O*NET OnLine: Occupational Information Network, Version 29.0.” https://www.onetonline.org/

Methodological References

[18] Hyndman, R.J. and Athanasopoulos, G. Forecasting: Principles and Practice. 3rd ed. OTexts, 2021. (SARIMA, Holt-Winters methodology)

[19] Hagberg, A.A., Schult, D.A., and Swart, P.J. “Exploring Network Structure, Dynamics, and Function using NetworkX.” Proceedings of the 7th Python in Science Conference (SciPy), 2008.

[20] U.S. Department of Justice and Federal Trade Commission. “Horizontal Merger Guidelines.” 2010 (revised 2023). (Herfindahl-Hirschman Index methodology)

[21] Mitchell, S., O’Sullivan, M., and Dunning, I. “PuLP: A Linear Programming Toolkit for Python.” https://coin-or.github.io/pulp/

[22] CHIPS and Science Act of 2022, Pub. L. No. 117-167, 136 Stat. 1366. (Semiconductor workforce development provisions used as analogous model)

The Institute for Homeland Security at Sam Houston State University is focused on building strategic partnerships between public and private organizations through education and applied research ventures in the critical infrastructure sectors of Transportation, Energy, Chemical, Water/Wastewater, Healthcare, and Public Health.

The Institute is a center for strategic thought with the goal of contributing to the security, resilience, and business continuity of these sectors from a Texas Homeland Security perspective. This is accomplished by facilitating collaboration activities, offering education programs, and conducting research to enhance the skills of practitioners specific to natural and human caused Homeland Security events.

[Institute for Homeland Security]

[Sam Houston State University]

Liu, Q. (2026). Predictive Analytics for U.S. Maritime Workforce Development and Supply Chain Resilience A Machine Learning Approach to Rebuilding Shipbuilding Capacity (Report No. 2026 -1042). The Sam Houston State University Institute for Homeland Security. https://doi.org/10.17605/OSF.IO/RE9KG

INSTITUTE FOR HOMELAND SECURITY
SAM HOUSTON STATE UNIVERSITY