# DATA SOURCES AND ATTRIBUTION

## Primary Employment & Wage Data

### Bureau of Labor Statistics (BLS)

- **Source** : U.S. Bureau of Labor Statistics - Occupational Employment and Wage Statistics (OEWS)
- **Dataset** : May 2024 National Occupational Employment and Wage Estimates
- **URL** : https://www.bls.gov/oes/current/oes_nat.htm
- **Specific Data Used** :
- Employment counts by occupation (SOC codes)
- Median annual wages by occupation
- Occupation titles and descriptions
- **Coverage** : ~830 detailed occupations representing the entire U.S. workforce
- **Data Retrieved** : November 2025
- **Reliability** : Official government statistics. Gold standard for U.S. employment data
- **Note** : Check for May 2025 OEWS release (typically published Q4 each year). Update employment counts and wage figures when available.

---

## Automation Risk & AI Impact Research

### 1. Frey & Osborne (2013)

- **Study** : "The Future of Employment: How Susceptible are Jobs to Computerization?"
- **Authors** : Carl Benedikt Frey and Michael A. Osborne, Oxford Martin School
- **URL** : https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf
- **Data Used** :
- Occupation-specific automation probability scores (0-100%)
- Task analysis (routine vs. non-routine)
- Engineering bottlenecks for automation
- **Methodology** : Machine learning classification of 702 occupations based on nine key variables
- **Key Finding** : 47% of U.S. employment at high risk of automation
- **Weight in Composite Score** : 20% (reduced from 40% -- study predates generative AI era and overweights traditional automation of physical/routine tasks)
- **Limitation** : Does not account for LLM-driven disruption of cognitive and creative work

### 2. Goldman Sachs Research (2023-2025)

- **Reports** :
- "The Potentially Large Effects of Artificial Intelligence on Economic Growth" (2023)
- "Generative AI and the Future of Work" (2023-2024)
- Annual AI impact updates (2024-2025)
- **URL** : https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html
- **Data Used** :
- Generative AI exposure scores by occupation
- Task-level automation potential
- Timeline estimates for AI capability development
- **Key Finding** : 300 million jobs globally could be affected by AI automation, with 25% of work tasks automatable by 2030
- **Weight in Composite Score** : Combined 30% with OpenAI/Eloundou (see #9)

### 3. McKinsey Global Institute (2024)

- **Study** : "Generative AI and the future of work in America"
- **URL** : https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america
- **Data Used** :
- Occupation automation timelines (early, mid, late adoption scenarios)
- Activity-level automation analysis
- Sector-specific displacement projections
- **Methodology** : Analysis of 2,000+ detailed work activities across 800+ occupations
- **Key Finding** : By 2030, activities accounting for up to 30% of hours worked could be automated
- **Weight in Composite Score** : 25% (increased from 20%)

### 4. World Economic Forum (WEF)

- **Report** : "Future of Jobs Report 2025" (5th edition, January 2025)
- **URL** : https://www.weforum.org/publications/the-future-of-jobs-report-2025/
- **Data Used** :

- Job displacement and creation projections
- Skills gap analysis
- Technology adoption timelines by industry
- Reskilling priorities and pathways
- **Methodology** : Survey of 1,000+ companies across 22 industry clusters and 55 economies
- **Key Findings** :
- 83 million jobs may be displaced while 69 million new jobs created by 2030 (net loss ~14 million, ~2% of employment)
- AI and big data skills ranked #1 fastest-growing skill
- Most declining roles: data entry clerks, administrative secretaries, cashiers, bookkeeping clerks
- Most growing roles: AI/ML specialists, data analysts, sustainability specialists, cybersecurity professionals
- **Note** : Updated from 2023 edition (which surveyed 803 companies across 45 economies with projections to 2027)

## 5. MIT Work of the Future Task Force (2024)

- **Study** : "The Work of the Future: Building Better Jobs in an Age of Intelligent Machines"
- **Authors** : MIT Task Force on the Work of the Future
- **URL** : https://workofthefuture.mit.edu/
- **Data Used** :
- Task complementarity analysis (human-AI collaboration)
- Industry transformation patterns
- Skill requirements evolution
- **Key Finding** : Emphasis on augmentation over replacement; middle-skill jobs most vulnerable

## 6. OECD Employment Outlook (2023-2024)

- **Report** : "OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market"
- **URL** : https://www.oecd.org/employment-outlook/
- **Data Used** :
- International comparative automation risk scores
- Policy implications and workforce transitions
- Cross-country employment trends
- **Key Finding** : 27% of jobs face high risk of automation across OECD countries

## 7. Brookings Institution Research

- **Study** : "Automation and Artificial Intelligence: How machines are affecting people and places" (2019-2024 updates)
- **Authors** : Mark Muro, Robert Maxim, and Jacob Whiton
- **URL** : https://www.brookings.edu/research/automation-and-artificial-intelligence/
- **Data Used** :
- Metropolitan area automation risk profiles
- Demographic vulnerability analysis
- Regional economic impacts
- **Key Finding** : Automation exposure varies significantly by geography and demographics; generative AI inverts traditional patterns -- higher-paid, white-collar workers now more exposed

## 8. Burning Glass Technologies / Lightcast

- **Data** : Labor market analytics and real-time job posting analysis
- **URL** : https://lightcast.io/
- **Data Used** :
- Emerging skill requirements
- Job transition pathways
- Real-time labor market demand signals
- Skill adjacency analysis

## 9. OpenAI -- LLM Exposure Analysis (2023) *NEW*

- **Study** : "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models"
- **Authors** : Tyna Eloundou, Sam Manning, Pamela Mishkin, Daniel Rock
- **URL** : https://arxiv.org/abs/2303.10130
- **Data Used** :
- Occupation-level LLM exposure scores
- Task-level assessment of GPT-4 capabilities
- Exposure ratings for ~1,000 occupations
- **Key Findings** :
- ~80% of U.S. workforce could have at least 10% of tasks affected by LLMs
- ~19% of workers may see 50%+ of tasks affected
- Higher-income occupations generally face greater exposure
- **Weight in Composite Score** : Combined 30% with Goldman Sachs generative AI data
- **Significance** : First rigorous LLM-specific exposure analysis; critical for recalibrating risk away from traditional automation models

## 10. IMF -- AI and the Future of Work (2024) *NEW*

- **Study** : "Gen-AI: Artificial Intelligence and the Future of Work" (IMF Staff Discussion Note SDN/2024/001)
- **Authors** : Mauro Cazzaniga et al.
- **URL** : https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379
- **Data Used** :
- Global AI exposure rates by country income level
- Inequality impact projections
- Policy response frameworks

- **Key Findings** :
- AI will affect ~40% of jobs globally
- In advanced economies, ~60% of jobs are exposed
- Of those exposed, roughly half may benefit (productivity gains), half may face displacement
- AI likely to worsen inequality within and between countries without policy action

### 11. Daron Acemoglu / MIT NBER (2024) *NEW*

- **Study** : "The Simple Macroeconomics of AI" (NBER Working Paper #32487)
- **Author** : Daron Acemoglu (MIT, Nobel laureate 2024)
- **URL** : https://www.nber.org/papers/w32487
- **Data Used** :
- Conservative macroeconomic modeling of AI productivity effects
- Task-level automation cost-benefit analysis
- **Key Finding** : AI will increase total factor productivity by only 0.53-0.66% over 10 years -- significantly more conservative than Goldman Sachs or McKinsey estimates
- **Significance** : Provides critical counterweight to bullish industry projections; used to temper extreme risk scores

### 12. Stanford HAI AI Index Report (2025) *NEW*

- **Report** : "Artificial Intelligence Index Report 2025"
- **URL** : https://aiindex.stanford.edu/report/
- **Data Used** :
- Annual AI capabilities benchmarking
- AI investment and adoption trends
- Labor market impact aggregation across multiple studies
- **Significance** : Most comprehensive annual compendium of AI data; cross-references findings from many sources listed above

### 13. Harvard Business School / BCG Field Experiment (2023) *NEW*

- **Study** : "Navigating the Jagged Technological Frontier"
- **Authors** : Fabrizio Dell'Acqua et al., with Boston Consulting Group
- **URL** : https://www.hbs.edu/ris/Publication%20Files/24-013_d9b45b68-9e74-42d6-a1c6-c72fb70c7571.pdf
- **Data Used** :
- Empirical productivity effects of AI tool usage by knowledge workers
- Task performance quality with/without AI
- **Key Findings** :
- Consultants using GPT-4 completed 12.2% more tasks, 25.1% faster, with 40% higher quality
- Performance DECREASED on tasks outside AI's capability frontier when workers over-relied on AI
- **Significance** : Best empirical evidence of actual AI productivity effects; informs reskilling recommendations

---

## Task & Skill Classification

### O*NET Database

- **Source** : U.S. Department of Labor O*NET (Occupational Information Network)
- **Version** : 23.1 (competency framework updated December 19, 2024)
- **URL** : https://www.onetonline.org/ | Database: https://www.onetcenter.org/database.html
- **Data Used** :
- Detailed task descriptions by occupation
- Skills, knowledge, and abilities requirements
- Work activities and context
- Technology skills required
- **Usage** : Task classification into routine/cognitive/creative/interpersonal categories
- **Coverage** : 974 occupations with detailed task-level data
- **Weight in Composite Score** : 25% (increased from 15%)

### World Economic Forum Skills Taxonomy

- **Source** : WEF Future of Jobs Report 2025 Skills Framework
- **Data Used** :
- Core competencies by occupation
- Emerging skills demand
- Skills stability and disruption ratings

---

## Technology Adoption & Diffusion

### International Federation of Robotics (IFR)

- **Report** : "World Robotics Report 2024"
- **URL** : https://ifr.org/
- **Data Used** :
- Industrial robot adoption rates by sector
- Automation technology penetration timelines

- **Key Finding** : Robot density increasing 5% annually; 4 million industrial robots by 2025

### Gartner Technology Hype Cycles

- **Report** : Annual Hype Cycle for Emerging Technologies
- **URL** : https://www.gartner.com/en/research/methodologies/gartner-hype-cycle
- **Data Used** :
- Technology maturity timelines
- Adoption curve estimates
- Plateau of productivity projections for AI/ML technologies

---

## Education & Training Cost Data

### College Board -- Trends in College Pricing 2025

- **Report** : "Trends in College Pricing and Student Aid 2025"
- **Published** : November 2025
- **Data Covers** : Academic Year 2025-26
- **URL** : https://research.collegeboard.org/trends/college-pricing
- **Data Used** :
- Published tuition and fees by institution type (2-year, 4-year public/private)
- Net price after financial aid
- Room and board estimates
- **Key Figures (AY 2025-26)** :
- Public 2-year in-district: $4,150/year
- Public 4-year in-state: $11,950/year
- Public 4-year out-of-state: $31,880/year
- Private nonprofit 4-year: $45,000/year
- **Usage** : Primary source for undergraduate education cost estimates in transition cost calculator

### Education Data Initiative (educationdata.org)

- **Source** : Urban Institute / Education Data Initiative
- **Updated** : 2025-26 academic year
- **URL** : https://educationdata.org/average-cost-of-college
- **Data Used** :
- Graduate/professional degree cost breakdowns
- Master's degree average total cost (~$62,280)
- Doctoral tuition by institution type
- Professional degree costs (MD, JD, PharmD, DDS, etc.)
- **Underlying Sources** : NCES, IPEDS, College Board
- **Usage** : Primary source for graduate and professional degree cost estimates

### NCES/IPEDS -- Integrated Postsecondary Education Data System

- **Source** : National Center for Education Statistics
- **URL** : https://nces.ed.gov/ipeds
- **Data Used** :
- Institutional-level tuition and fee data
- Financial aid statistics (percentage receiving aid, average award)
- Program completion rates and timelines
- **Most Recent Release** : January 2026 (Fall 2024 enrollment data)
- **Limitation** : Tuition data lags ~2 years; supplemented with College Board current-year data

### Course Report -- Coding Bootcamp Data (2026)

- **Source** : Course Report annual bootcamp market survey
- **URL** : https://www.coursereport.com/coding-bootcamp-ultimate-guide
- **Data Used** :
- Average bootcamp cost: ~$13,584
- Cost by specialization (software engineering ~$15K, data science ~$14K, cybersecurity ~$10.6K)
- Program duration (average 14 weeks)
- **Usage** : Cost estimates for technology certificate and bootcamp pathways

### Professional Program Cost Sources

- **Medical (MD/DO)** : AAMC (Association of American Medical Colleges) tuition survey data
- **Legal (JD)** : ABA (American Bar Association) law school tuition data
- **Pharmacy (PharmD)** : AACP (American Association of Colleges of Pharmacy)
- **Dental (DDS/DMD)** : ADA (American Dental Association) dental school data
- **Other Professional** : Respective accreditation body published tuition data
- **CDL Training** : FMCSA program cost data
- **Aviation** : FAA-certified flight school program costs

---

## Reskilling & Workforce Transition

### Coursera Skills Report

- **Report** : "Global Skills Report 2024"
- **URL** : https://www.coursera.org/skills-reports/
- **Data Used** :
- Popular reskilling pathways
- Skill proficiency benchmarks

- Learning duration estimates
- **Usage** : Realistic reskilling program costs and timelines

### LinkedIn Economic Graph Research

- **Data** : "Jobs on the Rise" and workforce transition analytics
- **URL** : https://economicgraph.linkedin.com/
- **Data Used** :
- Successful career transition pathways
- Skills transferability analysis
- Job-to-job mobility patterns
- **Usage** : Reskilling success rate estimates

### U.S. Department of Labor - CareerOneStop

- **Source** : Career exploration and training resources
- **URL** : https://www.careeronestop.org/
- **Data Used** :
- Training program costs
- Certification requirements
- Educational pathway timelines

### Association for Talent Development (ATD)

- **Data** : Corporate training cost benchmarks
- **Usage** : Enterprise reskilling cost estimates and ROI projections

---

## Economic & Wage Projections

### BLS Employment Projections (2023-2033)

- **Source** : Bureau of Labor Statistics Employment Projections Program
- **URL** : https://www.bls.gov/emp/
- **Data Used** :
- 10-year occupational growth projections
- Industry employment trends
- Labor force participation forecasts
- **Methodology** : Economic modeling and demographic analysis
- **Note** : Check for 2024-2034 update when available

### Federal Reserve Economic Data (FRED)

- **Source** : Federal Reserve Bank of St. Louis
- **URL** : https://fred.stlouisfed.org/
- **Data Used** :
- Wage growth trends
- Labor market indicators
- Economic productivity metrics

---

## Methodology Notes

### Automation Risk Score Calculation

The automation risk scores (0-100) in the WDF database represent a composite assessment derived from:

- **Frey-Osborne base probability** (20% weight) -- reduced from 40%; foundational but predates generative AI
- **Goldman Sachs + OpenAI/Eloundou generative AI exposure** (30% weight) -- increased; reflects LLM-specific disruption of knowledge work
- **McKinsey automation timeline** (25% weight) -- increased from 20%
- **Task routine-ness analysis from O*NET** (25% weight) -- increased from 15%

Scores are normalized to a 0-100 scale where: - **0-30** : Low automation risk (highly resistant to automation) - **31-60** : Moderate risk (partially automatable) - **61-100** : High risk (highly automatable in near to medium term)

**Key change (March 2026)** : Weights rebalanced to reflect the generative AI era. The original 40% Frey-Osborne weighting overemphasized traditional automation of physical and routine tasks. The new weighting better captures LLM-driven disruption of cognitive, creative, and white-collar work -- consistent with findings from OpenAI/Eloundou (80% of workers affected), IMF (60% in advanced economies), and Goldman Sachs (300M jobs globally).

### Task Breakdown Methodology

Task percentages (routine, cognitive, creative, interpersonal) are derived from: 1. O*NET work activities analysis (version 23.1) 2. WEF core competencies framework (2025 edition) 3. Expert synthesis of task characteristics

Percentages sum to 100% for each occupation.

### Technology Adoption Speed

Classification of adoption speed (slow/moderate/rapid) based on: 1. IFR robotics adoption data 2. Gartner technology maturity assessments 3. Historical technology diffusion patterns in similar sectors 4. Current AI capability advancement rates

### Displacement Projections

Displacement estimates combine: 1. Automation risk scores (using new composite weights) 2. Technology adoption timelines 3. Current employment levels (BLS data) 4. Sector-specific economic modeling

---

## Data Quality & Limitations

### Strengths

- Official government statistics (BLS) as employment/wage foundation
- Peer-reviewed research (Oxford, MIT, Harvard) for automation analysis
- Leading consulting firms (McKinsey, Goldman Sachs, BCG) for current AI impact assessments
- International organizations (WEF, OECD, IMF) for global context
- LLM-specific exposure data (OpenAI/Eloundou) for generative AI calibration
- Conservative academic counterweight (Acemoglu/NBER) to prevent over-estimation

### Limitations

- Automation predictions are probabilistic and subject to uncertainty
- Technology adoption rates can vary significantly by region, company size, and sector
- New job creation from AI (not fully captured) may offset displacement
- Regulatory, social, and economic factors not fully modeled
- Data reflects 2024-2026 knowledge; AI capabilities advancing rapidly
- Acemoglu's conservative estimates suggest actual disruption may be slower than composite scores imply

### Data Currency

- BLS employment/wage data: May 2024 (most recent official release)
- Automation research: 2023-2025 publications
- AI impact assessments: Ongoing updates as of April 2026
- WEF Future of Jobs: 2025 edition (January 2025)
- O*NET database: Version 23.1 (December 2024)
- Education costs: AY 2025-26 (College Board Trends 2025, Education Data Initiative)
- Bootcamp costs: 2026 (Course Report)
- Next BLS OEWS update: Check for May 2025 release
- Next College Board Trends update: November 2026

---

## Citation Guidelines

When using WDF data in reports, presentations, or publications, please cite:

**For Employment/Wage Data:** U.S. Bureau of Labor Statistics. (2024). Occupational Employment and Wage Statistics. Retrieved from https://www.bls.gov/oes/

**For Automation Risk Analysis:** Workforce Disruption Forecaster database compiled from: Frey, C. B., & Osborne, M. A. (2013); Goldman Sachs Research (2023-2025); Eloundou, T. et al. (2023); McKinsey Global Institute (2024); World Economic Forum (2025); IMF (2024); Acemoglu, D. (2024).

**For Comprehensive Citation:** Data sourced from Workforce Disruption Forecaster (WDF), integrating U.S. Bureau of Labor Statistics employment data with automation risk research from

Oxford University, McKinsey Global Institute, Goldman Sachs, OpenAI, IMF, Harvard Business School, World Economic Forum, and other authoritative sources (2024-2026).

---

## Contact & Updates

For questions about data sources, methodology, or to report data issues: - Review the complete methodology documentation - Check for updated research publications from key sources - Monitor BLS for annual OEWS data releases (typically May each year)

**Last Updated** : April 2026 **Next Review** : May 2026 (BLS data update cycle) / November 2026 (College Board Trends update)