ANDREW HARRIS

Vancouver, WA • linkedin.com/in/thatsgooddata • andrew@abharrismethods.com • 360.904.6802

PROFESSIONAL SUMMARY

Director of Web Data Acquisition and engineering leader with nine years at ZoomInfo, scaling the company from startup to public-market leader. Architect of billion-scale crawlers, agentic extraction pipelines, and vectorized data platforms powering commercial intelligence. Designed and scaled a 200M+ query/month SERP engine and distributed crawler monitoring 1B+ pages/month. Expert in distributed systems, real-time data ingestion, cloud-native orchestration, and high-precision LLM/SLM-based extraction. Proven builder of high-performance teams across backend, data infrastructure, and machine learning. Delivers scalable, fault-tolerant data platforms that drive product innovation and enterprise revenue.

CORE COMPETENCIES

Data Infrastructure: Postgres (pgvector), AWS (S3, EC2, Lambda, Kinesis), Redis, Docker, GitHub Actions, Apache (Airflow, Kafka, Spark, Beam), Snowflake, BigQuery, GCS, Varnish

Languages & Frameworks: Python, SQL, FastAPI, Flask, Scrapy, Playwright, Selenium, BeautifulSoup, Node.js, Neo4J, Spanner, NLTK, spaCy, Go, PyArrow, LlamaIndex, Groq, OpenTelemetry, RAG

Search & Crawling: SERP orchestration, proxy management, scale rate-limiting, diff-based crawling, Apache Nutch, Frontera, Crawl4AI, Colly, Guzzle, headless fingerprinting, CAPTCHA resolution

Machine Intelligence: OpenAI & Anthropic LLMs, custom SLMs, embeddings, entity resolution, enrichment models, evaluation heuristics, LangChain, LangGraph, HuggingFace, Vertex AI

Data Strategy: Vendor integration, acquisition signals, financial modeling, ingestion pipelines, agnostic ELT, CRM integration, social listening, data lineage, systems architecture, GDPR, CPPA, CPRA, PIPL

Leadership: Hiring & scaling, high-performance teams, technical mentorship, OKR design, reliability SLAs, CI/CD, cross-functional leadership (ML, DevOps, Product)

WORK HISTORY

ZoomInfo | *October 2019 - Present (2 yrs 9 mos)* | *Vancouver, WA*

Senior Manager, Agentic Web Data Acquisition (Mar 2023 - Present)

Senior Manager, Software Engineering (Apr 2021 - Mar 2023)

Manager of Data Strategy (Oct 2019 - Mar 2021)

- Architected and scaled a SERP engine processing 500M+ monthly searches and distributed crawler frameworks handling 1B+ pages/month with intelligent throttling, caching, and diffing
- Lead multiple teams of engineers across backend, DevOps, and ML; establish agile roadmaps, velocity metrics, and CI/CD standards
- Pioneered agentic extraction systems using LLMs and domain-specific SLMs for high-precision entity and event detection
- Integrated and normalized multi-format datasets from two global vendors (covering 600M+ people and 100M+ companies), building ETL pipelines and graph linking logic that expanded enrichment coverage and accuracy across ZoomInfo's platform.
- Developed news-velocity tracking and realtime acquisition-signal models linking information flow to public-market behavior
- Mentored 50+ engineers across data, backend, and automation teams, building culture of experimentation and distributed architecture

DiscoverOrg | *July 2016 - October 2019 (3 yrs 4 mos)* | *Vancouver, WA*

Operations Analyst (2018 - 2019)

Special Projects Analyst (2016 - 2018)

- Joined as one of first 150 employees; contributed to foundational systems enabling growth from startup to enterprise scale
- Authored voicemail fingerprinting algorithms validating phone-number ownership, ingesting tens of millions of data points
- Accelerated contact-growth from <1M to >100M verified entities through analytical models and data quality improvements
- Conducted specialized research projects supporting product development and data acquisition initiatives during critical growth phase

EDUCATION

Master of Public Affairs (M.P.A.), Applied Policy | *Washington State University*

Bachelor of Arts (B.A.), History & Development of Language | *Winthrop University* | *Sigma Tau Delta - Chapter President*

NOTABLE ACHIEVEMENTS

- Transformed ZoomInfo's acquisition codebase into a billion-page, enterprise-grade web intelligence platform driving core product data.
- Advanced from analyst to director scope assuming ownership of the company's mission-critical acquisition infrastructure and strategy.
- One of the first 150 employees instrumental in growing engineering and data infrastructure through 3,500+ headcount and an explosive IPO.