

Kenneth Flynn

Richmond, VA +1 (781) 424 8671 | ken@flynndata.com | [linkedin.com/in/kmflynn5](https://www.linkedin.com/in/kmflynn5)

Experience

Founder & Principal Data Engineer

Flynn Data Services LLC, Richmond, VA | February 2026 - present

- Founded boutique data consultancy serving B2B SaaS companies (\$5M–\$50M ARR) with complex MarTech stacks and limited in-house data engineering resources
- Deliver Data Stack Audits identifying cost optimization opportunities across cloud infrastructure, SaaS tooling, and data pipelines for Snowflake-based data environments
- Provide Fractional Data Engineering and CRM/RevOps pipeline architecture, offering on-demand senior expertise for companies without full-time data teams

Lead Data Science Engineer

Cybrary, Inc., Riverdale, MD | June 2024 - February 2026

- Reduced Algolia search costs by 92% annually by auditing actual search volume and right-sizing the contract from Enterprise to GROW tier, preserving full search performance for millions of catalog queries
- Engineered backend app-engagement data pipelines for Cybrary's Salesforce-to-HubSpot CRM migration, rebuilding data synchronization to eliminate middleware dependency and support the organization's 66% reduction in annual CRM licensing costs; deprecated Hightouch integration by building custom solution
- Engineered a custom Reverse ETL engine (Airflow + Cloud Run) syncing Snowflake data to HubSpot every 15 minutes, reducing CRM data latency by 96% vs. the prior 24-hour vendor sync; architected multi-modal ingestion layer supporting real-time event streaming, micro-batch enrichment, and configurable warehouse-to-CRM jobs
- Architected a high-velocity HubSpot Ingress Engine (Python/Cloud Run) processing 32k+ monthly user signups and 14k+ onboarding submissions with idempotent upsert logic, E.164 phone normalization, and real-time NeverBounce email validation (52k+ invalid contacts scrubbed at a 9.4% rate)
- Implemented GDPR/CCPA-compliant consent management via Google Tag Manager and iubenda (43 tags configured); restored attribution for 30-40% of

enterprise SSO signups by engineering custom UTM/Gclid header extraction in the PHP Auth API during SSO handshakes

- Introduced repository-wide pytest standards with dslib-powered GCP service mocks, shifting from manual production testing to automated CI gates that catch 3-5 critical regressions per month and enabling reliable Friday deploys

Senior Data Engineer

Living Security, Inc., Austin, TX | November 2022 - June 2024

- Designed and architected Living Security's Events API that allows customers to use their own infrastructure to send event data to the Unify platform and bypass the need for third-party access
- Lead the development of Unify's Behavior Score model to help summarize a user's risk in a more human-readable and customizable format
- Contributed to Unify's analytics risk scoring implementation to improve both flexibility and accuracy

Senior Data Engineer

Cybrary, Inc., Riverdale, MD | December 2018 - November 2022

- Architected Cybrary's first centralized reporting and data analytics system aggregating data for millions of users from a variety of sources
- Oversaw the migration of this system from AWS to GCP and its maturation over time to become more reliable and self-serve as the number of end users grew
- Built an ETL pipeline responsible for routing thousands of new and existing leads per day for marketing and sales campaigns in near real-time
- Lead a project to generate closed captions for Cybrary's 74,000 audio minutes of educational video content, and designed a scoring algorithm to target captions of lower quality to make the best use of more expensive human editors, increasing the average quality score by 29.74

Field Data Scientist

Forcepoint, Baltimore, MD | September 2017 - December 2018

- RedOwl Analytics was acquired by Forcepoint in 2017
- Joined the Professional Services group to help sell, implement, and support Forcepoint UEBA deployments utilizing a human-centric analytical approach to identify and mitigate insider risk

Field Data Scientist

RedOwl Analytics, Baltimore, MD | August 2015 - September 2017

- Worked with both regulatory surveillance and information security customers to design and implement analytical strategies to address insider risk use cases
- Developed robust ETL pipelines to enable both the broader data science and sales teams to pursue new analytical opportunities using a variety of novel data sources
- Prototyped solutions for analytical and ETL product gaps identified in the field

Fellow

Insight Data Science, New York, NY | August 2015 - June 2015

- Developed Python code to parse MySQL database queries and utilized a logistic regression classifier to predict user churn with 80% accuracy from recent mobile app usage logs

Postdoctoral / Graduate Researcher

University of New Hampshire, Durham, NH | September 2009 - February 2015

- Created a data pipeline utilizing tools such as Trimmomatic, bowtie2, and breseq to assist university researchers extract actionable insight from microbial genome sequencing data
- Developed an algorithm to reduce the dimensionality of metagenomic sequencing datasets up to 99.89% to direct limited resources to the most functionally important mechanisms

Skills

Languages: - Python (expert) - SQL (expert) - R (intermediate) - Perl (intermediate)

Tools: Snowflake, Airflow, Terraform, Docker, Fivetran, Snowplow, Hightouch, Looker, dbt Core, Apache NiFi, UNIX

AWS: Step Functions, Lambda, S3, SNS, SQS, API Gateway, Secrets Manager, IAM, Transcribe, RDS

GCP: Cloud Run, Cloud Functions, Pub/Sub, GCS, Secret Manager, Cloud Build

CRM/MarTech: Salesforce, HubSpot, Pardot, Marketo

Education

Ph.D. in Microbiology

University of New Hampshire, Durham, NH | 2014

B.A. in Biology

Colby College, Waterville, ME | 2009 - Minor in Japanese language

Fun Facts

- Seventh best speedrun time in the world for Mega Man II on Game Boy in the Any% category (<https://megamanleaderboards.net/index.php?game=21>)