

Luka Berg

☎ (212) 555-0184 ✉ luka.berg@example.com 💼 linkedin.com/example/lukaberg 📍 New York, NY 10011

SUMMARY

Lead Machine Learning Engineer with 15 years of software engineering experience and 13 years delivering **production machine learning systems** for high-volume content and audience platforms. Brings hands-on depth across **data collection**, data exploration, **feature engineering**, batch training, deep learning, statistical modeling, and **low-latency online serving**. Builds scalable AWS infrastructure, Spark and Databricks pipelines, REST microservices, and observable distributed systems that support recommendation, object detection, automated tagging, and retrieval experiences. Known for connecting product, editorial, data, and engineering goals with practical technical decisions. Eager to help Northstar Media Systems create useful personalized experiences while mentoring engineers, reducing delivery risk, and strengthening production ownership.

EXPERIENCE

Lead Machine Learning Engineer

January 2024 - Present

Brightline Digital Networks

New York, NY

Leads technical direction for a content intelligence platform spanning recommendation, automated tagging, retrieval, data pipelines, online inference, observability, and production operations. Guides engineers and partners with product, editorial, and **data science** leaders on initiatives that connect audience needs with scalable machine learning delivery.

- Set platform direction for recommendation, automated tagging, and retrieval across web, mobile, and connected devices.
- Rebuilt inference workflows with AWS **Step Functions**, Lambda, Glue, SQS, SNS, and **Personalize** for reliable recovery.
- Standardized Spark and Databricks **feature libraries**, aligning training data with low-latency REST serving.
- Guided design reviews, risk assessments, testing plans, and production readiness across product and editorial teams.
- Expanded Datadog metrics, logs, dashboards, and alerts for model freshness, latency, drift, and pipeline failures.
- Introduced **CI/CD**, model validation, rollback, **incident response**, and documentation practices across Scrum teams.

Principal Machine Learning Engineer

July 2020 - December 2023

HarborPoint Streaming Labs

Jersey City, NJ

Owned architecture for personalized content-ranking services supporting large streaming and news applications. Combined event-driven data platforms, **Spark** and Databricks machine learning pipelines, deep learning workflows, REST services, and team coaching to improve discovery and operational confidence.

- Owned architecture for personalized ranking services supporting millions of daily recommendation decisions across applications.
- Designed **REST** microservices with Kafka and **Kinesis** signals, preserving scalable boundaries for candidate generation.
- Led Spark and **Databricks** pipelines for feature preparation, **batch training**, evaluation, and model promotion.
- Delivered deep **learning** tagging workflows with editorial queues, improving classification and content discovery.
- Translated audience goals into experimentation plans with product managers, editors, platform engineers, and analysts.
- Mentored six engineers while standardizing code review, integration testing, CI/CD, and post-incident learning.

Senior Machine Learning Engineer

January 2017 - June 2020

CivicWave Media Technologies

Boston, MA

Developed production personalization services for regional publishers, connecting behavioral data, catalog metadata, and contextual signals with audience-facing products. Advanced **data exploration**, feature engineering, object detection, semantic tagging, quality controls, and Agile delivery while **coaching** junior engineers.

- Developed recommendation services combining behavioral data, metadata, and contextual signals for publisher experiences.
- Built Spark and Databricks workflows for exploration, **feature engineering**, training sets, and model evaluation.
- Implemented Python REST endpoints and cloud queues for low-latency ranked content retrieval.
- Prototyped **object detection** and semantic tagging for image and video libraries with editorial review.
- Added **Kafka** ingestion checks and data-quality controls before downstream model training workflows.
- Coached junior engineers on experiments, testing, service ownership, Scrum delivery, and **technical documentation**.

Machine Learning Engineer

August 2013 - December 2016

SignalForge Analytics

Chicago, IL

Built supervised learning products from audience behavior and content metadata, supporting recommendation, churn-risk, and targeted-distribution use cases. Established repeatable modeling, batch scoring, REST integration, monitoring, and stakeholder **communication** practices that prepared experiments for production.

- Converted audience behavior and content metadata into supervised datasets for recommendation and churn-risk models.
- Validated statistical and deep learning features through holdout testing and controlled product experiments.
- Created Spark batch scoring workflows with **AWS** services, improving repeatability for model releases.
- Developed REST APIs exposing ranked results with documented **contracts**, failures, and performance expectations.

- Established **monitoring** for data freshness, prediction volumes, and serving errors across production pipelines.
- Presented evaluation tradeoffs to product, analytics, and engineering partners for practical release decisions.

Software Engineer

May 2011 - July 2013

Redwood Software Group

Austin, TX

Built Java and Python services for distributed content and customer-data platforms, gaining foundational experience with REST APIs, relational data, asynchronous processing, testing, cloud deployment, and production support. Contributed within Scrum teams while progressing toward machine learning platform engineering.

- Built Java and Python services for distributed content platforms with REST APIs and asynchronous processing.
- Implemented ingestion and transformation components for analytics and downstream reporting workflows.
- Added unit and integration tests as new data sources and API consumers entered production.
- Configured queues, scheduled jobs, logs, and deployment scripts under senior engineering guidance.
- Investigated production defects through log analysis and reproducible test cases without disrupting customer services.
- Contributed to Scrum planning, peer reviews, design documentation, and technical demonstrations.

PROJECTS

Audience Personalization Platform 📅 2023

A ranking service became a daily product touchpoint. Built event-driven feature pipelines, model evaluation, and REST inference for timely content recommendations across streaming experiences.

Multimodal Content Tagging 📅 2020

Editors needed faster ways to classify growing media libraries. Combined object detection, deep learning inference, and review workflows for searchable image and video collections.

Retrieval-Augmented Content Discovery 📅 2025

Search results needed richer context. Prototyped RAG workflows that joined content metadata, semantic signals, and low-latency services for audience-focused discovery.

LEADERSHIP & AWARDS

- Engineering Excellence Award, HarborPoint Streaming Labs, 2022
- Technical Innovation Award, CivicWave Media Technologies, 2019
- Machine Learning Mentorship Lead, Brightline Digital Networks

EDUCATION

Bachelor's Degree in Computer Science

2011

The University of Texas at Austin GPA: 3.8

Austin, TX

***Coursework:** Machine Learning, Data Structures, Algorithms, Statistical Methods, Distributed Systems*

CERTIFICATIONS

- AWS Certified Machine Learning - Specialty 📅 2023
- Databricks Certified Machine Learning Associate 📅 2022

TECHNICAL SKILLS

- **AWS Services:** Step Functions, Lambda, Glue, SQS, SNS, Personalize
- **Big Data Platforms:** Spark, Databricks, Kafka, Kinesis
- **Machine Learning:** Deep learning, statistical modeling, recommendation systems
- **ML Operations:** Batch training, model evaluation, model promotion
- **Data Engineering:** Data collection, data exploration, feature engineering
- **Backend Development:** Python, Java, REST microservices
- **Serving Systems:** Low-latency serving, online inference, ranking services
- **Observability Tools:** Datadog, metrics, logging, monitoring
- **Delivery Practices:** CI/CD, automated testing, model validation
- **Reliability Practices:** Incident response, rollback procedures, recovery paths
- **Workflow Systems:** Workflow orchestration, data pipelines, feature libraries
- **Development Methods:** Agile, Scrum, design reviews

SKILLS

- Machine Learning Engineering
- Deep Learning
- Statistical Modeling
- Recommendation Systems
- Object Detection
- Automated Tagging
- RAGs
- Feature Engineering
- Data Quality
- Model Evaluation
- Batch Training
- Online Serving
- Python
- Java
- REST

PROFESSIONAL AFFILIATIONS

- Volunteer Workshop Mentor, New York Data Science Community
- Member, Association for Computing Machinery
- Member, Institute of Electrical and Electronics Engineers

LANGUAGES

- English (Native)
- German (Intermediate)

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST