

Luke Weber

@lkwbr

lkwbr.github.io

lkgwbr@gmail.com

BEWARE: THIS IS AN INCOMPLETE AND POTENTIALLY INACCURATE FIRST DRAFT

Experience

Omic, Chief Engineer (Feb 2020- Mar 2021)

- Created a system to support the identification and treatment of any 7,000+ human diseases in under 5 minutes; currently deployed internally to generate novel small molecules (drugs) to treat mantle cell lymphoma (a notoriously hard-to-treat breed of cancer).
- Architected and implemented said system as a multi-omic, highly-scalable, distributed, team-based web application—*Omic OS*—to support collaborative data science, healthcare, and bioinformatics workflows.
- Coded and collaborated with scientific team to build an end-to-end drug creation pipeline utilizing generative attention and convolution networks for selected proteins; generated hundreds of millions of compounds.
- Constructed a novel multi-million node biomedical knowledge graph and crafted graph-based ML models for the identification of "druggable" protein targets within disease networks; mined from nearly 30 million scientific articles using BioBERT along with proprietary, imputed intrinsic and external data.
- Co-developed an end-to-end Mutational MORTalityY (M-MORTY) patient predictor using 40k SARS-CoV-2 genomes with derived variants, and a drug pipeline for targeting the non-structural protein, NSP3 with around 0.8 AUROC; developed CAPTCHA-resistant crawler for obtaining the data.
- Co-designed and co-implemented the product UI/UX; conducted usability testing with scientists.

- Co-developed Genome And Medical Record DataBase (GAMER-DB) and tooling in integration with our knowledge graph; prototyped auto-ETL workflows for bringing customer data into GAMER.
- Invented an intent-based ecosystem within product consisting of apps, tools, and pipelines as a marketplace driven by internal NLP for identification of patient cohorts, analytical intent, and over 25k biomedical terms.
- Built on-prem Beowulf cluster "nucleus" for processing our most intensive computations.
- Co-led COVID-19 initiative c19.ai and our subsequent partnership with NSVD, the country's largest pool of scientific talent deployed to understand and treat COVID-19.
- Started and contributed to open-source healthcare record (C-CDA) crawler to enable citizen's with exposure to COVID-19 to anonymously contribute patient records.
- Led a fantastic and driven team of 15 total genomic and drug-discovery scientists, full-stack engineers, data scientists, and web designers.
- Conducted product demonstrations to key partners and clients with collective market caps of over 350 billion; co-gathered multiple "champions" and key decision makers in the process.

Omic, Research Scientist (Apr 2019 - Feb 2020)

- Co-developed personalized variant-based genetics pipelines for measurement of drug metabolism capability in addition to personalized health and wellness reports.
- Moved the company to stable test-production environments with rigorous tests on infrastructure and core science.
- Worked on theoretical feasibility of many systems, including: a reference-based genome compression and a knowledge graph-based search engine.
- Constructed state-machine and Alexa-based conversational assistant—our so-called *ConverState* Machine—for our medical practitioner-facing product, *Omic MD* (precursor to *Omic OS*).

- Built immunotherapy assessment pipeline for somatic cancer tissue genomes using "tumor mutation burden" (TMB).
- Prototyped patient "cost bloom" prediction, with near-70% accuracy.

Vizinet, Research Engineer (May 2016 - Apr 2019)

- Extended a grant-funded project *AIRPACT-Fire* (now *VIZINET*) by the Joint Fire Science Program as a central machine learning platform for in-field air quality analysis of wildfires.
- Led research and development on a research initiative for inferring particulate matter (PM2.5) concentration levels using convolution networks trained on data collected from crowdsourced smartphone images.
- Drafted and implemented entire website and mobile app UX from obscure and generic product requirements.
- Built dynamic Android app interface for submitting image data and observations, adaptable to new air-quality algorithms provided by environmental researchers.
- Designed Android app as client to server backend—submitting posts, authenticating users, recording server algorithm output, and managing local submissions.
- Built app management suite to sync with activity with application-, server-, and database-specific functionality, improving performance by ~3x on critical tasks.
- Performed backend maintenance, security, and upgrades, in addition to project documentation.
- Directed formal pre-production testing with ~15 academic, government, and layperson users.

Microsoft, Software Engineer, Contract (Dec 2017 – Sep 2018)

- Worked within Microsoft Azure, Digital Subject Rights (DSR) team, for processing petabytes of sensitive user data within their primary data stores.
- Owned external-facing service and optimized the processing of privacy commands by up to 10x.

- Moved team to more efficient command data store to reduce external dependencies.
- Developed team's I/O package for efficiently reading and writing 2M+ daily public requests to Azure storage.
- Built internal Azure service for tracking and processing thousands of user-sensitive data streams a second (32x faster than previously achieved) ensuring edge-case data is properly disposed.
- Built APIs used by over 1K Microsoft Service Teams (e.g., Office, Xbox, Windows, Skype) for validating exports and deletes on their sensitive user telemetry.
- Constructed entire team's test infrastructure and validation service used within our code development life-cycle.
- Coded Cosmos DB scripts for weekly hard-deletes on many terabytes of user telemetry.
- Acted efficiently to resolve high-severity incidents directly with Service Teams, PMs, and senior staff at the onset of GDPR.
- Generated thorough internal developer documentation and health monitors for owned services.

Washington State University, Research Assistant (Jan 2017 – May 2017)

- Worked on theory of ML-based task assignment system *AgiPal* within human organizations.
- Designed as enhancement to SCRUM work-item assignment using machine learning.
- Went from abstract idea to concrete, functional implementation in approximately 4 months.
- Built library for extracting qualitative developer experience from any given Git repository.
- Read and summarized 10+ computer science papers for supporting my research.
- Pursued research under Dr. Jana Doppa and Dr. Venera Arnaoudova.

Central Washington University, Research Assistant, Pro bono (Jan 2013 - Mar 2013)

- Developed Java Swing desktop application (for Mac, Windows, Linux) to help learning-disabled students with reading 50+ preprocessed textbooks as well as pasted text.
- Utilized PDF-to-text conversion to simplify content by near $\frac{2}{5}$ (best approx.) and with under 5% trial error.
- Implemented text-to-speech functionality.

Education

Causal Inference, Columbia University (completing Apr 2021)

Computational Neuroscience, University of Washington (completing Apr 2021)

Differential Equations, Massachusetts Institute of Technology (completing Jun 2021)

Deep Reinforcement Learning Nanodegree, Udacity (completed Feb 2019)

- Multi-agent, deep RL; actor-critic methods and Deep Deterministic Policy Gradients; policy gradient methods and Trust Region Policy Optimization (TRPO) and Proximal Policy Optimization (PPO) and Generalized Advantage Estimation (GAE); policy-based methods and REINFORCE

Flying Car and Autonomous Flight Engineer, Udacity (completed Aug 2018)

- Developed drone 3D motion planned solutions.
- Wrote non-linear cascading controller for Unity-simulated quadrotor.
- Created drone attitude and position estimator using Extended Kalman Filters (EKFs).
- Fixed-wing aircraft autopilot and control problems.

Deep Learning Specialization, [deeplearning.ai](https://www.deeplearning.ai) (completed Jan 2018)

- Sequence Models and CNNs.
- Hyperparameter Tuning, Regularization, and Optimization.

- Built NNs from scratch. Projects areas from Neural Style Transfer to self-driving cars.

B.S. Computer Science, Washington State University (graduated May 2017)

- Honors: Cum Laude (GPA 3.7/4.0) and President's Honor Roll (6x).
- Relevant coursework: Machine Learning (graduate-level), Structured Prediction (graduate-level), Artificial Intelligence, Operating Systems, SDE I/II, and Databases.

Awards and Certifications

Google Code Jam 2021 — *In progress*

Crimson Code 2017 — *Finalist* — _____

Crimson Code 2015 — *Finalist* — Created aggregated conversational AI (similar but in no way competitive to the modern-day Google Assistant).

MTA: Software Development Fundamentals (2012)

MTA: Networking Fundamentals (2012)

MTA: Web Development Fundamentals (2012)

Patents, Essays, and Publications

Coming soon.

Products and Projects

Omic OS (2021)

A computational biology and AI platform currently being used to identify novel targets and generate drugs in oncology.

Webercoin (2021)

A Haskell-based cryptocurrency based off the Bitcoin whitepaper, implementing core blockchain and networking; used in the Weber family for transacting chores.

Vizinet (2021)

An end-to-end data collection and data science platform; inferring particulate matter concentration from low-grade smartphone Images.

A Very Weber Christmas (2018)

Pygame-based, Entity-component-system (ECS) game where each Weber family member has unique powers in their quest to defeat Santa (or each other).

Passive, portable metabolism tracker (2017)

Portable metabolic tracking app connected via BLE protocol to CO2 sensor with microchip, targeting users with type-1 and type-2 diabetes. Achieved about 90% accuracy for body fuel-source estimation under stable (inactive) conditions.

Automated task assignment (2016)

Prototyped automated task assignment system within human organizations, with applications found in SCRUM workflows.

Fake news classifier (2016)

Fake news classifier as Chrome extension and AWS backend, built for Facebook. Used Bayesian learning methods from article URLs as features, deriving additional training data through crowdsourcing user feedback. Surpassed 95% testing accuracy (via k-fold cross-validation).

Newtonian physics simulator (2015)

Newtonian physics emulation engine for interacting and planning object trajectories with ___ in plain JavaScript.

Winter sport weather tracker (2012)

Java-based (JSP) web application for parsing and aggregating raw government sensor data and displaying it in a clean interface.

Hobbies and Miscellaneous

- Ran Irving Marathon participant in 3:51:00 and helped raise money to help fight childhood obesity.
- Ran Mount Si in 2:05:00'.
- Ran Manashtash Ridge in 33:04'.
- Released 26 experimental tracks on Soundcloud.

- Released 8 experimental short films on YouTube.
- Learning German; estimated Duolingo fluency: beginner.
- Chess ELO score is 880 on 5+0 Blitz.