

Ryan N. Lichtenwalter

CONTACT INFORMATION

Denver, Colorado, USA

E-mail: rlichtenwalter@gmail.com

Web: lichtenwalter.site

GitHub: github.com/rlichtenwalter

RESEARCH INTERESTS

Network science, link prediction, machine learning, large data sets, distributed computation, data stream processing, and engineering education.

EDUCATION

The University of Notre Dame, Notre Dame, Indiana, USA

Ph.D., [Computer Science](#), May 2012

- Thesis Topic: *Link Prediction*
- Adviser: [Nitesh V. Chawla](#)
- Area of Study: Data Mining and Network Science

M.S., [Computer Science](#), May 2009

- Thesis Topic: *Knowledge Acquisition in Dynamic Data*
- Advisor: [Nitesh V. Chawla](#)
- Area of Study: Data Mining

B.S., [Computer Science](#), May 2006

- *cum Laude*, With Honors in Engineering

REFEREED JOURNAL PUBLICATIONS

Meloto C.B., Slade G.D., **Lichtenwalter R.N.**, Bair E., Rathnayaka N., Diatchenko L., Greenspan J.D., Maixner W., Fillingim R.B., and Ohrbach R. [Clinical predictors of persistent temporomandibular disorder in people with first-onset temporomandibular disorder: a prospective case-control study](#). In: *Journal of the American Dental Association*, 2019.

Zorina-Lichtenwalter K., **Lichtenwalter R.N.**, Zaykin D.V., Parisien M., Gravel S., Bortsov A., and Diatchenko L. [A study in scarlet: MC1R as the main predictor of red hair and exemplar of the flip-flop effect](#). In: *Human Molecular Genetics*, 2019.

Meloto C.B., Benavides R., **Lichtenwalter R.N.**, Wen X., Tugarinov N., Zorina-Lichtenwalter K., Chabot-Doré A.J., Piltonen M.H., Cattaneo S., Verma V., Klares R. 3rd, Khoury S., Parisien M., and Diatchenko L. [Human pain genetics database: a resource dedicated to human pain genetics research](#). In: *Pain*, 2018.

Yang Y., **Lichtenwalter R.N.**, and Chawla N.V. [Evaluating link prediction methods](#). In: *Knowledge and Information Systems*, 2015.

Lichtenwalter R.N. and Chawla N.V. [Vertex collocation profiles: theory, computation, and results](#). In: *Springer Plus*, 2014.

Ercsey-Ravasz M., **Lichtenwalter R.N.**, Chawla N.V., and Toroczkai Z. [Range-limited centrality measures in complex networks](#). In: *Physical Review E*, 2012.

Davis D., **Lichtenwalter R.**, and Chawla N.V. [Supervised methods for multi-relational link prediction](#). In: *Social Network Analysis and Mining*, 2012.

Lichtenwalter R.N. and Chawla N.V. [LPmade: Link prediction made easy](#). In: *Journal of Machine Learning Research*, 2011.

Lichtenwalter R.N., Lichtenwalter K., and Chawla N.V. [A machine learning approach to autonomous music composition](#). In: *Journal of Intelligent Systems*, 2010.

- Lichtenwalter R.N.**, Zorina-Lichtenwalter K., and Diatchenko L. [Genotypic data in relational databases: efficient storage and rapid retrieval](#). In: *Proceedings of the 21st European Conference on Advances in Databases and Information Systems*, 2017.
- Lichtenwalter R.N.** and Chawla N.V. [Link prediction: fair and effective evaluation](#). In: *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 2012.
- Lichtenwalter R.N.** and Chawla N.V. [Vertex collocation profiles: Subgraph counting for link analysis and prediction](#). In: *Proceedings of the International Conference on World Wide Web*, 2012.
- Lichtenwalter R.N.** and Chawla N.V. [DisNet: A framework for distributed graph computation](#). In: *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 2011.
- Davis D., **Lichtenwalter R.**, and Chawla N.V. [Multi-relational link prediction in heterogeneous information networks](#). In: *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 2011.
- Lichtenwalter R.N.**, Lussier J.T., and Chawla N.V. [New perspectives and methods in link prediction](#). In: *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, July 2010.
- Lichtenwalter R.N.**, Lichtenwalter K., and Chawla N.V. [Applying learning algorithms to music generation](#). In: *Proceedings of the Indian International Conference on Artificial Intelligence*, December 2009.
- Lichtenwalter R.N.** and Chawla N.V. [Adaptive methods for classification in arbitrarily imbalanced and drifting data streams](#). In: *Lecture Notes in Computer Science: New Frontiers in Applied Data Mining (PAKDD 2009 Workshops)*, 2010.
- Chawla N.V., Thain D., **Lichtenwalter R.N.**, and Cieslak D.A. [Data mining on the grid for the grid](#). In: *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing*, April 2008.

- Financial Industry Regulatory Authority (FINRA)**, Denver, Colorado, USA
- Director of Technology – Enterprise GenAI Solutions* **May 2026 to Present**
Engineering Training
- **Enterprise GenAI Training:** As curriculum designer and sole content developer, built FINRA’s corporate GenAI curriculum in five months with Kiro-accelerated development: 6 tracks, 20 courses, and 87 modules, with 617 slides, more than 190 code sandbox exercises, 83 interactive widgets, and 695 concept-check questions, tailored to FINRA’s mindset, toolset, and skillset. Topics span prompt engineering, retrieval techniques, agents and MCP, evaluation, production LLM concerns, and GenAI for leaders.
Technologies: Kiro, MCP, agents, RAG, DeepEval
- Director of Technology – Regulatory GenAI* **January 2025 to May 2026**
- **Platform Turnaround:** Assumed leadership of FILLIP, FINRA’s regulatory GenAI platform, as a research-grade application in production and re-engineered it into an enterprise product while it served live users. Within a year it was the organization’s most widely adopted platform, with more than 1,800 weekly active users, over half of all employees.
 - **Delivery and Reliability:** Instituted CI/CD, automated testing, blue-green deployment, and test hygiene, moving from ad hoc to weekly releases and eliminating recurring production outages.

- **Architecture:** Defined the platform architecture: an LLM gateway (LiteLLM) over Amazon Bedrock, a programmatic API at parity with the UI, and agentic retrieval (Google ADK, Strands) spanning LLM map-reduce over long and many documents, agent-generated queries over structured data, and RAG.
 - **Hands-On Engineering:** Wrote high-leverage core code the whole team built on: components of the LLM architecture, cross-stack structured logging and error handling, a unit-testable system prompt framework, core team libraries and tooling, and CI, setting high team standards for code quality and production engineering.
 - **Leadership:** Hand-picked and led a team of eight, hiring two. With the Senior Director, aligned Product Management, Operations, Legal, and the Enterprise Data Privacy Office on human-in-the-loop controls and user training.
 - **People Development:** Identified and invested in high-potential staff on the data science and FILLIP teams, overseeing five promotions, including a skip-level promotion, and recommended the successor who now leads the platform as Director.
- Technologies:** Python, Amazon Bedrock, LiteLLM, Google ADK, Strands, RAG, DeepEval, CI/CD, blue-green deployment

Director of Technology – Data Science **September 2022 to January 2025**

- **Center of Excellence:** Led a team of seven as FINRA's natural language data science center of excellence, delivering for business and technology partners and mentoring the team in setting technical direction as the work moved from classical NLP to large language models.
- **Regulatory Workflow Delivery:** Delivered summarization, extraction, and classification solutions across regulatory workflows, including regulatory summaries, reports, and SEC filing processing, saving thousands of hours.
- **Innovation:** Technical lead of the [2024 FINRA Createathon](#) team that won the hackathon's grand prize and its challenge division among 32 teams from 11 departments with NITRO, an LLM-assisted regulatory document-tagging prototype advancing FINRA's machine-readable rulebook.
- **Origination and Team Building:** Conceived the seed of FINRA's flagship GenAI platform, FILLIP, to demonstrate LLM capabilities quickly and drive business results. Hired three team members and managed a multi-million-dollar budget.

[Transamerica](#), Denver, Colorado, USA

Manager Data Science **March 2021 to August 2022**

- **Team Leadership:** Maintained a weekly one-on-one cadence, effective and productive weekly team meetings, and cross-team collaboration meetings. Overcame previously existing cross-team challenges, crafting an atmosphere of effective collaboration.
- **Project Portfolio Leadership:** Project portfolio included, as a sample, sales forecasting, advisor and customer segmentation, organizational network analysis, underwriting automation evaluation, claims fraud detection, and call center recording and transcription. Provided meaningful business-oriented guidance and technical leadership to team members across the project portfolio.
- **Hiring and Onboarding:** Successfully hired and onboarded three excellent full-time employees, four strong interns, and a rotation program participant in a challenging market thus maintaining team size and project output.
- **Stakeholder Management and Faithful Delivery:** Reset expectations in previously challenging relationships with business stakeholders through a record of responsiveness, thoroughness, and promptness both in personal interactions

and through allocation of team resources and prioritization to accomplish repeated on-time delivery of insights.

- **Platform Migration:** Managed data science and data engineering resources to accomplish planning, migration, training, and project transition of the data science team and a second analytics team from Domino Data Lab to Amazon SageMaker in service of cost savings. Provided several critical individual contributions, including custom compliant images and S3 project synchronization scripts. **Skills:** project management, SageMaker, ECR, EC2, S3, docker, bash

Lead Data Scientist

December 2019 to March 2021

- **Project Leadership:** Served as technical lead and mentor on two major project efforts: underwriting automation, leading a team of four while performing formal project management, managing contract resources, designing and implementing data science solutions, and interfacing with business stakeholders, technology partners, and engineering resources; and sales leads generation, mentoring a colleague in development of a sophisticated tiered model architecture consisting of geospatial modeling, PU learning for propensity modeling, and bipartite classification. **Skills:** project management, feature engineering, model evaluation, model deployment, software engineering
- **Full-Stack Data Science Contributions:** Across many projects, worked with business partners to identify problems, led technical efforts to drive identification of data science solutions, and implemented fast, effective, and business-aware queries, statistical test procedures, and models. **Skills:** Python, SQL, pandas, scikit-learn, hypothesis testing, automation, feature engineering, model development, model evaluation, business interfacing
- **Core Platform Enhancement:** Resolved issues and improved usability of analytics platform through innovative technical solutions drawing on breadth of background. **Skills:** docker, bash, Python, pip, conda, cryptography, PKI
- **Core Library Development:** Implemented major components that increased efficiency, capabilities, and compliance in daily work for the data science team, including a wrapper API with data preparation capabilities for entity resolution submissions; cryptographic extensions (with automatic key management) to pandas series, data frames, and functions; high-performance (far exceeding geopandas) geospatial indexing in C++ with Python bindings; a C++ implementation of the POPC clustering algorithm with Python bindings; `fast_read_sql` pandas interfacing for large SQL result sets with drastically (5+ times) reduced time and memory requirements versus pyodbc; Python bindings for my previously existing high-performance C++ implementation of the mRMR feature selection algorithm; a sophisticated model monitoring and cost monitoring framework with statistical and distributional introspection; a type inference and conversion library; a centralized code authentication and authorization library; and a configuration-driven SMB Python wrapper API. **Skills:** Python, pandas, C++, pybind11, bash, cryptography, SQL, feature selection, SMB, PostgreSQL, Amazon Redshift

Human Pain Genetics Lab, McGill University, Montreal, Quebec, Canada

Researcher

January 2014 to December 2019

- **Modeling and Machine Learning:** Developed and applied high-performance machine learning strategies and software to perform multivariate genetic analysis. **Skills:** data cleaning, feature selection, clustering, classification, wrappers, ensembles, decision trees, neural networks, nested cross-validation, model selection, model evaluation, data visualization, graph analysis, C, C++, R, caret, WEKA, Cytoscape

- **Statistical Analysis:** Performed a variety of statistical analyses on TB-scale data sets and acted as statistical advisor to other researchers. *Skills:* R, awk, Perl, bash, linear and logistic regression, PCA, mRMR, SKAT, correlation analyses, entropy analyses, variable transforms, gnuplot
- **Database Administration and Development:** Developed sophisticated and efficient database-driven web portals and utilities to allow collaborators to perform data aggregation and analyses. *Skills:* PostgreSQL, MySQL/MariaDB, SQL, C, stored procedures, 5NF, indexing, partitioning, optimization, bash, Perl, R, GNU make, git, bit manipulation, credential management, role-based authorization
- **Full-stack Web Development:** Designed and implemented various web sites and single-page applications to manage clinical studies, facilitate and track purchase orders, and improve lab public relations. *Skills:* HTML 5, Javascript, CSS, JQuery, JQuery UI, Twitter Bootstrap, PHP, GNU make, R, Perl, SQL, stored procedures, TLS PKI, Apache, nginx, cron, git, L^AT_EX, geolocation
- **System and Network Administration:** Deployed and administrated secure computational and network infrastructure to manage and analyze data sets in GB and TB scales. *Skills:* Enterprise Linux, Debian, Windows Server, managed UPSes and PDUs, managed switches, Infiniband, multipath SAS topologies, Cisco IOS, pfSense, iptables, cron, RAID, LVM, KVM/QEMU virtualization, LDAP, Kerberos, NFS, SMB, SGE, RH Kickstart, preseed, cron, bash, SELinux, ZFS, OpenVPN
- **Technical Liaising:** Acted as universal technical liaison with internal and external entities. *Skills:* purchasing, information security, secure data exchange, contractor evaluation, offsite training, grant reviews, instrument support, systems integration, L^AT_EX

The University of Notre Dame, Notre Dame, Indiana USA

Post-doctoral Researcher

May 2012 to August 2012

- **Graph Analysis:** Developed mathematical models and algorithms for large, complex network data. *Skills:* RDF/triplestore, graph visualization, link prediction, semantic network analysis, C, C++, SGE
- Worked in collaboration with [Battelle Memorial Institute](#) and [Pacific Northwest National Laboratory](#) (PNNL).

Lockheed Martin Corporation, Eagan, Minnesota USA

ELDP – Software Development

September 2006 to July 2007

- Rotated as part of the Lockheed Martin Engineering Leadership Development Position, a position for high-potential new hires.
- **Software Architecture:** Performed software design, prototype development, and testing for a high-visibility fledgling project. *Skills:* Java, JSP, CSS, Javascript, Oracle Database, SQL, geolocation, bash
- **Software Development:** Extended large mission-critical legacy systems. Developed software prototypes under tight time constraints, receiving spot performance award. *Skills:* C, C++, Ada, Qt
- **Database Development:** Assisted in Oracle Database schema design to support software development efforts. *Skills:* Oracle Database, SQL

ELDP – Configuration Management

June 2006 to September 2006

- Rotated as part of the Lockheed Martin Engineering Leadership Development Position, a position for high-potential new hires.
- **Configuration Management:** Managed build architecture and code for large mission-critical code bases and performed extensive task automation. *Skills:* bash

OPEN SOURCE SOFTWARE

Independent Projects

Author

- [dir2text](#) Python library and CLI, published on PyPI, that renders directory trees and file contents in LLM-ready formats with streaming output and token counting
- [kdtree](#) cache-efficient C++ k-d tree with an STL-style container-adaptor interface
- [popc](#) C++20 implementation of Powered Outer Probabilistic Clustering for binary feature data

Human Pain Genetics Lab, McGill University

Author

- [PostgreSQL Genomics](#) relational database genotypic data representation
- [various](#) efficient PostgreSQL user-defined types implemented in C
- [mRMR](#) feature ranking software by Peng et al. (substantially improved implementation)

The University of Notre Dame

Author

- [VCP](#) link analysis software
- [LPmade](#) link prediction software (published in *Journal of Machine Learning Research*)
- [WEKA](#) data mining software ([SMOTE](#) component)

Contributor

- [WEKA](#) data mining software

Further projects at github.com/rlichtenwalter, plus dozens of private projects across GenAI, machine learning, systems programming, and infrastructure, developed in a home lab with a self-hosted Gitea server; glad to discuss any in depth.

PRESENTATIONS

Financial Industry Regulatory Authority (FINRA)

- More than two dozen internal presentations on generative AI, FINRA's regulatory GenAI platform, and best practices, including multiple presentations to the entire FINRA Technology organization, 2022–Present

TEACHING EXPERIENCE

The University of Notre Dame, Notre Dame, Indiana USA

Sole Instructor and Lecturer

January 2013 to May 2013

- Instructor for CSE 66725: Directed Readings in Machine Learning and Computer Vision
 - Spring 2013 (1 section)
 - Responsible for one semester of 3-hour guided reading group meetings for a graduate-level three-credit course.
 - Developed syllabus, selected readings, and offered insights and experience.

Sole Instructor and Lecturer

August 2010 to December 2010

- Instructor for CSE 30331: Data Structures
 - Autumn 2010 (1 section)
 - Responsible for one semester of 75-minute lectures for a required junior-level three-credit course.
 - Developed syllabus, assignments, course project, and exams.

Graduate Teaching Assistant

January 2009 to May 2009

- Assistant for CSE 40768/60768: Network Phenomena
 - Spring 2009 (1 section)
 - Assisted Network Phenomena instructor.
 - Provided technical support and web maintenance services.
 - Graded course assignments.
 - Assisted in the preparation of exam materials.

Guest Lecturer

August 2008 to December 2010

- Lectured for CSE 20211: Fundamentals of Computing I
 - Autumn 2008 (1 section), Autumn 2009 (1 section), Autumn 2010 (1 section)
 - Responsible for writing and presenting 50-minute lectures for a required sophomore-level three-credit course.

Undergraduate Mentor

August 2009 to March 2010

- Student: Jake T. Lussier
 - Served as research advisor to and published refereed paper with student.

AWARDS

[Financial Industry Regulatory Authority \(FINRA\)](#)

- [Createathon](#) grand prize and first place in the Data Drag Reduction System challenge division (3000 USD), as technical lead of an 11-person team, 2024
- Ten BOND Awards from FINRA's peer- and manager-nominated recognition program, six with monetary awards, 2022–Present

[The University of Notre Dame](#)

- [Center for Research Computing](#) Award for Computational Sciences and Visualization (1000 USD), 2012
- [Arthur J. Schmitt Presidential Fellowship in Science and Engineering](#) Fellow (> 140,000 USD), 2007–2011
- [National Science Foundation Graduate Research Fellowship Program](#) Honorable Mention, 2009

[Lockheed Martin Corporation](#)

- Spot Award (250 USD), 2007
- Service Award, 2007

[The Timken Company](#)

- Timken Company Scholar ($\approx$ 140,000 USD), 2002–2006

[National Merit Scholarship Corporation](#)

- National Merit Scholar (2500 USD), 2002