Karl D. Gierach

123 Regal, Irvine, CA 92620

(949) 300-1525 email: karl@gierach.net

http://www.linkedin.com/in/karlgierach

POSITION STATEMENT

Self-driven individual with excellent communication skills, innovative machine learning ideas, and experience in data mining and information retrieval, who is eager to provide both leadership and an individual contribution within a machine learning, data science, or AI group.

KEY STRENGTHS

Possesses a wide variety of deep technical skills, focused on highly stable, large-scale data analytics pipelines utilizing machine learning / statistical algorithms. Incorporates visualizations of data and algorithmic results, to communicate benefits to stakeholders. Experience working with and contributing to open source projects.

Experience in leadership and management, working with internal and external team members, including engineers, managers, stakeholders, other scientists, patent attorneys. Experienced with managing people at an individual level; weekly one-on-ones, annual reviews, and goal setting. Well versed in product development skills such as industry research, keeping up with industry news, and proposing product solutions to stakeholders.

WORK EXPERIENCE

Principle Data Scientist, INNOVID, Marina Del Rey, CA [12/19 - present]

- Leading the US-based Data Science Team within the Measurement Group, providing machine learning and statistical-learning solutions for Connected TV, Linear TV and Omni-Channel Digital Marketing.
- Perform research, pitching, storytelling, proof-of-concepts, model design & evaluation, while collaborating to produce production code that solves business problems.
- Authored a reach and frequency estimation pipeline for a major Linear TV partner, providing the firm the technology basis to compete with incumbent players such as Nielsen.
- Produced CTV reach and frequency estimation methods coupled with highly flexible per-publisher reach optimization. Led the team to define and implement multiple Extrapolation methods for Reach, Frequency, Overlap, Unique Reach, and Demographics.
- Initiated the build-out of an internal Household Graph product for enhanced accuracy of Linear & CTV measurement.
- Leading numerous strategic initiatives which are currently works in progress.
- Helped hire & build our first offshore remote data science team. Collaborated with UK-based DS team for bi-directional technology transfer post-merger with TV-Squared.

Principle Data Scientist, CONVERSION LOGIC, Santa Monica, CA [03/18 – 12/19]

• Bringing my expertise in Data Science to the firm's attribution (measurement) solution within the context of Digital Marketing and Advertising technology. Applying ensemble models comprised of Gradient Boosted Trees, FTRL, and Neural Networks to provide enterprise grade attribution measurement. Investigating causal methods for Marketing Science. Investigating hybrid models combining ecological-type data with individual data for attribution using privacy-sensitive platforms.

Head of Data Science, CAKE MARKETING, Newport Beach, CA [07/15 - 12/18]

• Spearheaded the Data Science initiative within the firm, from the ground up, focusing on developing Big Data analytics pipelines in a cloud computing context, for attribution (measurement), predictive, and prescriptive analytics within the context of Digital Marketing and Advertising technology.

• Perform Data Science Product R&D by finding and reviewing relevant academic / journal publications, industry reports, and competitive intelligence.

• Directed the team comprised of Data Scientists and QA Engineers, delegating projects and assignments.

• Architect and present the Data Science Roadmap, illustrating benefits of Data Science to the firm.

• Perform hands-on implementation of advanced machine learning algorithms, exercised with web traffic data originating from multiple sources / media types (e.g. Social Media, Display, Email, etc).

• Architected, Coded and Optimized the Big Data / Analytics Pipeline in Apache Spark / Hadoop, which leveraged a variety of data source types (SQL, Bulk Storage, etc).

• Performed thought leadership by authoring (2) U.S. Patents in the applied machine learning and statistics space, along with publicly facing Blogs (http://techblog.getcake.com/) describing these techniques.

• Communicate results with graphs / visualizations illustrating data characteristics and algorithmic results using "R" (ggplot2) and Python (Seaborn), D3, as well as Domo (a Tableau-like product).
• Actively inspired the team to experiment with modern data science techniques including various forms of Regression, Bayesian Inference, Factorization Machines, Causality, Time Series Analysis, Frequent Pattern Mining, Hidden Markov Models, Matrix Factorization for Recommendation, Pregel, Neural Networks with TensorFlow (Deep Learning), Linear Optimization, among others.
• Utilized NLP technology, namely Word2Vec and WordNet as a metric for referrer URL clustering.
• Proposed analytics supporting software architectures including Data Management Platforms (DMP), and Customer Data Platforms (CDP) enabling leverage of data across customers.

Principle & Lead Software Engineer, SOURCE THOUGHT, Aliso Viejo, CA [05/13 - 07/15]

• Researched, designed, and authored product's Big Data Discovery pipeline that performs deep data profiling, data similarity analysis, & advanced sampling techniques published in the data mining literature to process data at the Terabyte scale utilizing the Hadoop Ecosystem.
• Utilized NLP technology for feature engineering; n-grams, similarity algorithms, and edit distance algorithms for dataset comparison. Utilized hashing techniques to reduce run-time complexity.
• Used "**R**", to investigate the use of several classification algorithms to predict data relationships from features computed during the profiling phase. These algorithms include logistic regression, SVMs, boosted trees. Results visualized with GGPLOT "R" package. One ML technique boosted the Data Discovery F-measure from .6 to .92 when tested on a combined corpus of data sets.
• Architected & authored a Gateway service which performs native Hadoop integration services including HDFS, Map/Reduce, Pig, Hive, HCatalog, Oozie, and Impala. Architected and implemented a framework to allow product integration to any Hadoop Distribution (e.g. Cloudera, Hortonworks). Continuously prototype & investigate big-data technologies. Architected & implemented extensible data import framework. Continuously provide significant technical guidance & leadership to team members. The core product is written in HTML5, Node.js, Backbone, & MongoDB, with backend Java interfacing directly to Hadoop's services. Leverage modern technologies such as GitHub, Maven, Jenkins, and Pivotal in day-to-day operations.

President & Founder, VOICE IDENTITY, INC. Irvine, CA [03/09 – Present]

• Founder / CEO of Company, perform technical architecture, software development, and corporate vision for Cloud-computing based security software.
• Secured key partnership and licensing agreements.
• Presented company to various local investor groups, including an audience of over 500 people at the Irvine Chamber of Commerce.
• Handle all purchasing, advertising, proposals, and intellectual property initiatives.
• Conceived the flagship software product; a Voice Biometrics based Identity Management solution, designed to secure individual's digital identity over the internet that leverages the popular Open ID protocol, GWT, Hibernate, JDBC, etc.
• Performed custom development for primary client's voice user interface, Social Media (Twitter/ Facebook), Email, and SIP integration work using C#/.NET/Java focusing on Microsoft's Web and Speech platforms.

Principal Software Engineer, SPIREON, INC. Irvine, CA [11/11 – 04/13]

• Architected and coded Big Data frontend for an IOT/SAAS product designed to reliably handle tens-of-millions of events per day with high reliability.
• Served as initial MongoDB (Big Data) Architect interfacing with 10Gen (MongoDB parent company) to determine sharding, replication, and operations strategy.
• Formally designed and implemented the network facing and message parsing components in a horizontally scalable architecture using Java 1.7. Integrates with F5 load balancer, enterprise grade monitoring. Worked on the next generation product written in Groovy / Grails.

Senior Software Engineer / Architect, ORTIVA WIRELESS, INC. San Diego, CA [03/10 – 11/11]

• Served as Software Engineer and Architect on a real-time video optimization product positioned for wireless carriers, which optimizes bandwidth consumption while preserving high quality video using deep packet inspection.
• Managing and developing the implementation and testing cycle of Level 2/3/4 network processing code using Linux (Netfilter Queue) and RMIOS/NetLogic.
• Co-Inventor of U.S.A. patent titled "Data Path Processing". Embedded Linux based platform, designed for high availability, runs on ATCA telephony grade data center hardware.
• Performed integrations with 3rd party products (e.g. Sandvine) in the network processing space.

Senior Software Engineer / Architect, AVST, INC. Foothill Ranch, CA [11/00 - 03/09]

• Served as a Technical Leader and Architect and full life-cycle implementer for a speech enabled, unified messaging software product, and designed to serve enterprises of all sizes.
• Successfully mentored other programmers and managed several projects.

Key Roles Served / Accomplishments:

• Provided significant innovation and worked with IP attorneys to file multiple awarded patents.
• Fully managed the process of integrating with 3rd party vendors, while helping those vendors isolate defects in their products. Worked with customer support to diagnose field problems.

Key Technology Accomplishments:

• Invented adaptive statistical based audio end-pointing algorithms (Patented in USA).
• Performed expert-level multi-threaded, real-time, event driven, Java, C++, & C# development.
• Continuously created and prototyped advanced application features such as "whisper" (call waiting), "click to call", and high density, multi-node audio channel management.

Consultant, PHARMACIA, INC. Skokie, IL [08/00 - 11/00]

• Designed and implemented networking software for virtual environments using Java, Java3D, and Jini technologies. The network software ties together VRML based virtual worlds presented to users in different locations, allowing them to interact remotely and collaborate within them. Work was affiliated with the University of Illinois position noted below.

Visiting Research Engineer UNIVERSITY OF ILLINOIS. Chicago, IL [08/99 - 07/00]

• Worked as a visiting research engineer in the Industrial Virtual Reality Institute under Dr. Prashant Banerjee.
• Investigated the use of virtual reality in manufacturing. Introduced the results of the investigation in several lectures within a graduate level course. Coded the concepts investigated. Produced papers summarizing the research.

EDUCATION

UNIVERSITY OF ILLINOIS AT CHICAGO

M.S., Computer Science. Master's project titled "An Application of Genetic Algorithms to the Construction of Contour Approximations". Coursework focused on the fields of machine learning, computer vision, and distributed systems. Produced an in-depth, software-based course project in each graduate-level course. G.P.A. is 4.7/5.0.

UNIVERSITY OF ILLINOIS AT CHICAGO

B.S., Computer Engineering. Graduated 12/95. Cumulative G.P.A. is 3.95/5.0.

PUBLICATIONS

- Inventor, US Patent Pending Application, titled "Method for Targeting Electronic Advertising by Data Encoding and Prediction for Sequential Data Machine Learning Models".
- Inventor, US Patent Pending Application, titled "Non-converting Publisher Attribution Weighting and Analytics Server and Method".
- Inventor, US Patent 7,092,515, titled "VC-to-DTMF interfacing system and method".
- Inventor, US Patent 7,822,192, titled "Sound event processing with echo analysis".
- Inventor, US Patent 7,756,709 & 8,370,144, titled "Detection of voice inactivity within a sound stream".
- Inventor, US Patent 8,072,909 , titled "Apparatus and method for notification of a party in a telephone conference"
- Inventor, US Patent 7,992,196 , titled "Apparatus and method for performing hosted and secure identity authentication using biometric voice verification over a digital network medium"
- "Managing Hardware Devices and Software Objects in a Java Environment", Society of Mechanical Engineers / CASA Blue Book Series, 2001.
- Inventor*:* U.S. Patent Application 20120284373. DATA PATH PROCESSING.

TECHNICAL KNOWLEDGE SUMMARY

- **Machine Learning:** Studied and implemented several machine learning and optimization techniques including decision trees, neural networks, Naive Bayes, Hidden Markov models, reinforcement learning, and genetic algorithms, computer vision algorithms, audio analysis algorithms. Experience developing models with "**R**" (RStudio) using Regression, Logistic Regression, Decision Trees, Boosting, Bagging, Empirical Bayes, PageRank, Dimensionality Reduction and others.

- **Data Mining:** Applied use of Jaccard Similarity and similar coefficients as applied to N-Grams, Flajolet-Martin algorithm, Kelsey's Sim-Hashing algorithm for finding similar items. Use of Murmur Hash for advanced sampling techniques. Levenshtein distance algorithm. Experience with Market/Basket Analysis (Frequent Patterns), Stream Sampling, Locality Sensitive Hashing.
- **Hadoop:** Experience coding in the Hadoop ecosystem at many different levels including native Map/Reduce with mappers/reducers/combiners. Spark, Pig, Hive, Hive-Serdes, HDFS, Oozie, and other related technologies. Hands-on experience using Cloudera, Hortonworks, etc. Recognized as an Apache Hive open source contributor.
- **Programming Languages:** Scala, R, Java, Python, Javascript (Node.js), C++, C, Perl, Bash.
- **Database:** Snowflake, Athena, Redshift, Aurora, MongoDB, MySQL, SQL Server, Informix/Illustra, Exp. in DB design, stored procedures, replication, sharding.
- **Web / Service Development:** AJAX/RESTful Endpoints, CGI, Servelets, SOAP, Code Signing, ActiveX Controls, and Java Applets, Tomcat, Apache, IIS. Integrate to Facebook, Twitter, and other Social Media APIs.
- **Network Programming:** Expert with TCP/IP/UDP protocols, among others. Wireshark.
- **Advanced Programming Techniques:** Multi-processing/multi-threaded applications using Unix IPC methods, POSIX Threads.
- **Methods:** Object-oriented software development and methodologies including Agile, UML, and Carnegie Mellon University's PSP (Personal Software Process), and PMP.