

Bruce B. Campbell
(919)-808-8178
wavescholar@gmail.com

SUMMARY: Director of Data Science / Senior Data Scientist with over fifteen years of experience encompassing all phases of software and algorithm development, from requirements and design to coding, debugging, testing, documentation, user training, and maintenance. An optimist with a focus on customer support capable of solving complex design and implementation issues. Works well independently, but enjoys a collaborative team environment. Impeccable business skills developed in environments ranging from DOW30 to startup. Strong skills in leading the development of high performance/high availability analytic software systems. Experience writing reentrant code for shared memory multi-core environments. Experience with software development in an FDA regulated environment, and have personally managed a successful 510K application. Experience with systems level architecture design, team leadership, and management of offshore development teams.

EDUCATION:

University of Rochester: BS Physics 1991

University of Rochester: MS Applied Math 1992

University of Rochester: 2 additional years towards PhD in Applied Mathematics 992 to 1994.

North Carolina State University: MS Statistics June 2018.

CORE COMPETENCY: Implementation, commercialization, and integration of all varieties of business logic, image processing, computer vision, and machine learning algorithms into distributed computing environments. Ten years experience image processing, color science, and machine vision algorithms. Three years of experience in document classification, natural language processing, and ECM.

Languages: R, Python, C++, Java, Scala, SQL, C, Fortran, Matlab

Big Data Tools: AWS, Hadoop, Hive, Spark, MongoDB, Redis/Memcached

EXPERIENCE:

Sensus / Xylem (NYSE: XYL), Director of Data Science January 2016- May 2019 Curated and developed a rapidly growing data science department. Responsible for hiring and technical development of data science personnel, identifying and developing use cases for machine learning and data science, and developing and maintaining a strong statistics practice for internal analytics. Established a solid foundation of best practices for data science. Performed modeling and development for mission critical statistics and analytics using Spark, Hive, R. Responsible for maintaining academic collaborations at NCSU, Duke and developed a large sponsored machine learning research initiative at MIT (450,000\$ budget). Led deep learning R&D efforts in multiple domains.

Sensus, USA, Principal Data Scientist September 2014-2016 Responsible for establishing data science program and implementing data lake on a PB of time series data. First project resulted in 135 million dollar savings. Developed ingest workflow and Apache Hive data store ingesting more message per day than Twitter generates. Developed statistical models and machine learning analytics in Apache Spark using R, Scala, and Python. Developed AWS version of a data lake for data science activities resulting in lowered operational costs and increased stability. Developed numerous traditional machine learning applications such as anomaly detection, prediction, and clustering for commercial analytics software. Established academic collaborations at Duke University and North Carolina State University.

Cernostics Inc., Director of Informatics, 2011- 2014 Team lead and algorithm developer for machine learning and computer vision software for whole slide digital pathology analysis platform. Worked closely with pathologists and cancer immunologists to model oncogenesis. Developed academic collaborations with machine learning and image processing groups at Duke University, Carnegie Mellon University and The University of Pittsburgh. Developed object segmentation algorithms, and higher order graph diffusion algorithms for classification of very large tissue

images. Core machine learning algorithms deployed in a software framework developed in house that is fully reentrant, multithreaded, and highly optimized for speed, memory performance, and IO latency. Work resulted in multiple publications, and successful CLIA certified deployment of working cancer prognostic.

DynaVox Inc. (NASD:TOBII) , Senior NLP Engineer, 2010- 2011 Responsible for Natural Language Processing framework deployed in educational and alternative communication products. Tools were integrated into event-driven WinForm applications. Assisted in the development of language content database utilizing LINQ to SQL. Improved quality of word sense model using WordNet database. Implemented high-performance splay-tree data structure for storing n-gram frequency data. Utilized Microsoft Patterns and Practices Library in several projects.

Omnyx LLC (NYSE GE / UPMC), Senior Machine Vision Engineer, 2009-2010 Developed high performance distributed image processing framework and computer vision algorithms for digital pathology software in a .NET MVVM architecture. Lead architect and developer for certificate-based WCF security solution that was successfully deployed at multiple hospitals and research centers. Participated in the design and implementation of a pettabyte scale distributed storage system. Participated in the commercialization and integration of proprietary wavelet-based file format. Performed color calibration experiments.

Wavescholar Consulting LLC, Predictive Analytics Consultant, 2009 Commercialized graph spectra segmentation & clustering algorithms. Algorithms were prototyped in Matlab and commercialized in C++. Ported Arpack to Intel Fortran 90, Arpack++, SDPA, and METIS libraries to VS9 with Intel BLAS. Participated in various Imaging, Web, and ECM development projects.

Xerox Corp. (NYSE XRX), Senior Software Engineer 2008- 2009 Deployed and developed document and image classification technologies in a high volume distributed scanning application. Prototyped GPU based image processing and novel document classification algorithms. Resolved many software bugs related to memory management, image processing, .NET interop, and software configuration. Evaluated and used multiple OCR technologies. Performance benchmarked and characterized numerous imaging and document classification algorithms. Played lead role in application architect for a large contract with India Tax Department. Assisted in sale of job, and exceeded automation throughput targets. Efforts resulted in seven digit revenue gain.

Wavescholar Consulting LLC, Predictive Analytics Consultant, 2007-2008 Developed a high performance system for distributed financial analytics on basket ETFs & their derivatives using Townsend Analytics' .NET toolkit. Technologies for data and workflow management included SQL Server, MSMQ, ADO .NET, OLEDB DTC, and a service based architecture for deploying analytics. Technologies for analytics include wrappers for advanced memory management such as private Win32 heaps, AWE memory management, and interfaces to Intel IPP/MKL Libraries, Matlab, & Quantlib. Designed and implemented a multithreaded graph based algorithm workflow with smart pointer based results caching. Numerous NaN friendly, robust algorithms have been deployed on this framework. Real-time capabilities successfully demonstrated on the components of the SPY [500 dimensions] and IWM [2000 dimensions] simultaneously. Commercialized several machine vision applications prototyped in Matlab.

TissuelInformatics Inc., Icoria Inc. (NASD: ICOR) / Clinical Data Inc. (NASD: CLDA) / Bioimagine Inc, Senior Machine Vision Developer 2004-2007 Developed pixel, object, tile, and slide level models for very large (2-4GB) biological images in the field of digital pathology. Performed hierarchical statistical analysis on data sets up to 10TB in size. Developed standard and proprietary machine learning and signal processing algorithms at all scales of analysis. Pixel and object algorithms were written in C++/Java and deployed to a CORBA/Oracle distributed computing platform. Tile and slide models were developed in Matlab and R. Acted as principal investigator on numerous disease models. Reduced expert training load in inter-grade scoring models based on pair-wise comparative vision experiments using machine learning techniques. Statistical results were incorporated into NIEHS and NIH studies.

LIPAM International Inc. Data Analyst, 2004-2008 Provided data analysis, statistical analysis, and GIS informatics in the domains of microfinance and mass privatization. Results were published in multiple professional journals and delivered to organizations like the World Bank and The Gates Foundation.

Eastman Kodak Co. (NYSE: EK) Senior Software Engineer 1999-2004 Matured a commercial distributed image processing framework to CMM Level 3. Integrated and wrote complex image processing algorithms bringing the latest algorithms from R&D into existing software products. Solved difficult integration and optimization issues. Development efforts were cross platform on Win32, Solaris, and Linux and involved heavy use of OOP, C++, STL, and software patterns. Received recognition four times for rapidly solving critical customer issues. Redesigned core framework structure and integrated into existing code base. Aggressively sought out and exposed design and implementation issues before and after software release. Effectively communicated complex image processing, design, and integration issues to management and customers.

Awards & Recognition:

Eastman Kodak Image Science Career Development Program 2003-2004

Petroleum Research Council Fellow 1993-1994

Whittaker Prize in Mathematics 1992

Societies: American Statistical Association, SIAM, IEEE Information Theory Society

Other: National Outdoor Leadership School Graduate. Active in rock climbing and mountaineering