

# Jeffery L. Painter, JD | CV

301 Fayetteville St. Unit 2301 – Raleigh, NC 27601

📞 (919) 533-9024 • 📞 (919) 533-9024 • ✉️ jeff@jivecast.com  
🌐 jivecast.com • Data Wrangler

*"You can have data without information, but you cannot have information without data." - DKM*

## Summary

---

Innovative and creative technology leader with a history of building and leading game-changing processes and projects to empower life sciences with the best technologies. Track record of delivering global software solutions, leveraging proven technical, project management and people leadership/relationship management skills.

Leader in analytics, ontologies, and software design seeking opportunities to pioneer innovation. Over 15 years experience in leveraging and advancing the use of AI methods and machine learning in the drug informatics arena. Some highlights of my career in AI methods and technologies are noted below:

- *Knowledge Reasoning & Semantic technologies*  
Expert knowledge and developer of NLP methods, ontology and advanced knowledge representation. I have built and developed text processing methods using R, Python, C#/.Net and Java. Experienced with most of the popular open source text NLP packages including OpenNLP, Stanford NLP, R tm and TextBlob/NLTK in Python
- *Machine Learning*  
Led the development and implementation of a large scale predictive modeling engine for identifying patient sub-populations (random forest and hierarchical clustering) with varying levels of risk for the NC Medicaid/Medicare system (2014-2015 Patent pending). Tools and techniques applied have utilized SAS, R, tensorflow, DL4j (Java deep learning framework), WEKA data mining platform and SciKitLearn. Latest research has involved in-depth work with PPMI (Parkinson's Progression Markers Initiative) in the discovery of sub-populations with increased genetic risk burden for patient's with Parkinson's disease.
- *Automatic Classification & Information Retrieval*  
Expert panel member and invited guest speaker for the 2017 DIA meeting "People Will Talk" on using Bayesian classifiers for the automatic classification and identification of potential adverse events in social media for pharmacovigilance. Built various text analytics classifiers utilizing word2vec, Latent dirichlet allocation (LDA), and tf-idf transforms. Wrote and presented advantages of various classifications methods (a priori, multinomial Naïve Bayes, non-negative matrix factorization, SVM, k-nearest neighbors, etc.) of medical text at the July 2016 Joint Statistics Meeting (JSM)

- *Biotech Startup*

Since Feb-2019, I have been helping to build out a new biotech startup under the Flagship Pioneering venture labs (FL60). I was recruited as employee number 2, and in the past year, we have built a new lab from scratch, hired a full biology discovery team, and recruited some of the top talent in iPSC stem cell research and single-cell RNA sequencing analysis in the world. I have been directly involved in the entire interviewing and recruiting process, establishing our cloud based compute platform, and exposed to all the intricacies required in building a full-scale drug-discovery platform from scratch.

## Experience

---

### **Apache Software Foundation**

*Contributor and recent Member*

*Aug 2004–Present*

Apache Turbine and Apache Torque

- *Member* - Elected as a Member (<https://www.apache.org/foundation/members.html>) of the Apache Software Foundation Mar-2019
- *Nomination* - Nominated as a candidate for the ASF Board of Directors for 2020
- *Apache Turbine* - Long time contributor to the Apache Turbine project. The first OpenSource web framework which established a number of well-known subprojects (Apache velocity, Maven, Torque)

### **Flagship Lab 60**

*Senior Principal Scientist, Computational Biology and Data Science*

**Cambridge, MA**

*Feb 2019–Present*

Head of bio-informatics

- *Business Development* - Working with our lead statistical geneticist and business partners, in our first two months, worked to established computational methodology and biological platform for new drug-discovery model - Series A raising \$28 million (USD).
- *Data Imputation* - Developed new methods for handling missing value problems in longitudinal data by creating both Markov Models to discover population based trends for categorical data, and also employing the R package kmlShape (<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150738>) for the discovery of before unknown patient sub-populations using longitudinal trajectory analysis with kml (K-Means for Longitudinal Data ).
- *AWS* - All Flagship Pioneering labs are cloud first computing environments. Became quickly engrained in setting up, and administering our AWS infrastructure for analysis of WGS (whole-genome sequencing) data and the day-to-day needs of our computational group. My prior work in system administration and deep knowledge of Linux/UNIX environments lowered the threshold on my personal transition to a cloud based computing environment. Familiar with AWS pricing models, setting up EC2 instances, S3 buckets, EFS shared storage and networking (e.g. Route53, Private VPN and VPN gateway configuration)

**JiveCast**

CEO And Founder

Analytics and Consulting

**Raleigh, NC**

Sep 2016–Present

- *GWAS-MeSH Mapping* - Subcontractor to Deerfield (<https://deerfield.com>) : Developed new methods to map the GWAS catalog to MeSH headings in order to support proximity alignment between genes and medical conditions using both Lin and Resnik scoring algorithms through the UMLS-similarity Perl module.
- *Gen-AERS* - Subcontractor to GlaxoSmithKline Statistical Genetics group : Built genetically enhanced version of the FDA Adverse Event reporting database for use in developing biomarker targets using advanced AI methods for ontology alignment and establishing robust multi-axial ontological relations between drugs, adverse events and genetic indications using the UMLS and DrugBank.ca knowledge sets. Fully implemented all statistical measures in Java per the EudraVigilance recommendations (<https://bit.ly/2ZRFIha>) for screening potential signals from noise for target gene-AE identification.
- *Deep Learning Methods Development* - Project with Anexinet - Assisted in developing Recurrent Neural Nets to assist in the automatic identification of Adverse Events in call center data. Implementation developed in Java using DeepLearning4j framework.
- *Drug to Indication Maps* - Project with HAPi, LLC to develop a drug to indication mapping using custom built Java source and open knowledge sources (UMLS, etc).
- *Social Media for Pharmacovigilance* - Subcontractor to GlaxoSmithKline : Provide expert consulting and programming for delivering key insights from social media analytics. Projects included crowdsourcing, Java development for UI and data analysis. Provided development of various classifiers using OpenNLP, WEKA and tensorflow for the automatic classification and tagging of hundreds of thousands of social media posts for various projects

**GlaxoSmithKline**

Director Medical Advanced Analytics - US Medical Affairs

Deployed Innovations:

**Research Triangle Park, NC**

2006–2016

- *Medical Insight Explorer* - Web application and text analytics framework developed to help automate text classification and clustering for topic and trend analysis on unstructured voice of the customer feedback for US Medical Affairs. Enabling text analytics for a non-technical user base of 150 employees.
- *Insight Explorer for Social Media* - Work flow management system for the annotation of social media data for safety and pharmacovigilance - over 50,000 social media posts reviewed by the GSK safety listening laboratory leading to an improved understanding of the use, abuse, and benefits of marketed drugs.
- C#/.Net Predictive Modeling Engine deployed in CareTriage<sup>®</sup> which executes patient risk models in real time using random forest, LASSO, and inverse logit regression. Currently active in 2 state

medicare management systems.

<https://www.communitycarenc.org/media/files/care-triage-announcement.pdf>

- *CodeSlinger* - Biomedical terminology code management tool for epidemiology and observational data analytics.
- *BRAT* - Benefit Risk Analysis Toolkit - Collaboration between GSK, J&J, and AstraZeneca to standardize statistical analytical methods across the industry.
- Constructed the foundational terminology mappings enabling SafetyWorks and OMOP to support disparate data sources for unified statistical analysis.  
<http://www.ubc.com/tags/safetyworks>, <http://omop.org>

### **Kiasoft, Inc**

*Founder*

**Cary, NC**

*2001–2004*

- Developed Order Manager, a web based fulfillment and inventory management software system for LogosDirect.com.
  - Enabled real time inventory management and implemented predictive models to determine inventory requirements.
  - Deployed in 2003 and still in use, having completed over 100,000 orders.
- Elected as an Apache community contributor to the Turbine top level project.

### **MCI Worldcom**

*System Administrator*

**Cary, NC**

*2000–2001*

- Supported 150+ software developers who were responsible for global network provisioning services. Administered Sun Solaris, AIX, DB/2 and Oracle systems.
- Administrator for the Sun E10000 servers running WebLogic J2EE platform and Oracle databases.

### **SmartStart NC**

*Programmer*

**Raleigh, NC**

*1999–2000*

- Developed administration system for statewide financial managers written in Perl & PostgreSQL.

### **American Social Health Association**

*Programmer*

**RTP, NC**

*1999*

- Supported phone, network and servers for various CDC funded call centers.

### **UNC-TV**

*System Administrator*

**RTP, NC**

*1998*

- Supported phone, network and servers for television production and administrative use.

### **Osborne Computers**

*Systems Integrator*

**Dalton, GA**

*1996-1997*

- Administered BSD web servers, RADIUS dial-up service.
- Built custom PC desktop and networks for local clients.

## Video Presentations

---

I am not afraid to stand up in a room of a thousand strangers and give a talk...

- Jan 15, 2020 - "Why Apache? Trillions and Trillions Served: the documentary on The Apache Software Foundation"  
Video: <https://www.youtube.com/watch?v=YM5dLvNatRs>  
Jeffery begins at 0:25s
- Sept 2018 - ApacheCon 2018 - Lightning talks  
Video: <https://www.youtube.com/watch?v=FTR0CvmhXko>  
Jeffery begins 17:10-22:10, Title: "Apache Turbine isn't dead..."
- June 18, 2017, Drug Informatics Association (DIA) - 2017 Chicago II - "People Will Talk: Gathering Insights from Digital Listening" - A panel discussion on social media with leaders from FDA and EMA  
Program Details: <https://bit.ly/3fkSdbx>  
Video: <https://www.youtube.com/watch?v=mvzoIehe1EU> - 1hr 27min  
Jeffery speaks from 17:45 - 33:15
- 2014 - GSK/CCNC Collaboration interview  
Video: <https://www.youtube.com/watch?v=mjyqiGHR91M>

## Professional Development

---

- Building and Solving Optimization Models with SAS/OR™
- Data Mining with WEKA, Offered by University of Waikato
- Vaadin Fundamentals and Vaadin Advanced training
- First Line Leader course for new managers at GlaxoSmithKline
- Sun Certified administrator for Solaris 2.6/8.x
- Oracle DBA training track

## Awards

---

**2015:** Global Medical Excellence Award - Gold Winner

- *"Voice of the Patient: Project CRAWL - Contextualization of Real World Drug Use Through Social Listening"* The aim of this project is to deliver an automated, streamlined capability for using social media to enhance pharmacovigilance activities. The team delivered on several new Artificial Intelligence based capability enhancements to our safety organization, bringing the voice of the patient to the company.

**2015:** Global Medical Excellence Award - US Finalist

- *"Voice of the Patient: Project Insight Explorer - Fueling faster medical insights through the use of text analytics."* The team led this innovative solution to improve synthesis and delivery of high quality insight for Voice of the Patient by replacing manual data processing with an automated process using artificial intelligence and advanced NLP/text analytics. This effort provided more

robust and reliable insights in a timely manner to the matrix teams.

**2015:** GSK Bronze Award

- o *"Jeff's ability to work across boundaries to bring Insight Explorer to Canada will have significant impact on how Canada analyzes mVoC data. Without hesitation, Jeff stepped up to meet our needs and helped us navigate the IT process so we can utilize Medical Insight Explorer." – Matthew Swash*

**2013:** Exceptional Science Award

- o *"For developing computational tools for converting observational data for retrospective studies in Safety and Health Outcomes Research utilizing ontologies and high performance computing. Jeffery Painter is a key, highly regarded scientist in the area of high performance computing in observational data and has led and delivered projects for multiple teams throughout R&D." – Alan Menius*

**2011:** Exceptional Science Award

- o The Exceptional Science Award is given for outstanding contributions to research projects internally at GlaxoSmithKline. I was granted this award for my research into advancing the capabilities of automated ontology mapping through developing the use of a semantic measure of proximity.

## Patents

---

**Aug 12, 2015:** Application Filed

- o Electronically Predicting Corrective Options Based on a Sensed Physiological Characteristic.  
Inventors: Troy Trygstad, Jeffery L. Painter, Jr., and Alan Menius

## Education

---

**NCCU School of Law**

*JD, Concentration in Intellectual Property*

**Durham, NC**

*2011–2015*

**NCSU**

*Bachelor of Science, Applied Mathematics, Cum Laude*

**Raleigh, NC**

*2007–2009*

**NCSU**

*Bachelor of Science, Computer Science, Cum Laude*

**Raleigh, NC**

*2005–2007*

*Statistics Minor & Cognitive Science Minor*

## Technology Skills

---

- o Primary Languages: Java (Lambdas and Streams), Python, L<sup>A</sup>T<sub>E</sub>X, XML/HTML, JavaScript
- o Secondary Languages: Perl, C/C++, x86 assembly
- o Statistics & Data Mining: DL4j, WEKA, R, JMP, SAS (EM, Visual Analytics, Text Miner)
- o Databases: MySQL, PostgreSQL, MS SQL Server, Oracle, Teradata, Apache HBase, Berkeley DB
- o Frameworks: Apache Turbine, Vaadin, Spring, d3.js, jQuery
- o Containers: Apache Tomcat, jetty, WebLogic
- o Build Tools: Maven, Ant, Eclipse, JRebel, JUnit, SVN, Git, Jenkins
- o NLP: R tm, Python Textblob, Python nltk, Java OpenNLP, UMLS MetaThesaurus & MetaMap
- o Preferred environment: Ubuntu Linux, bash shell and Eclipse
- o Legal tools: FindLaw, WestLaw Next and Lexis/Nexis

## Publications

---

Publications available for review at <http://javastats.com/about.html>

1. Gartland, A., Bate, A., Painter, J.L. et al. *Developing Crowdsourced Training Data Sets for Pharmacovigilance Intelligent Automation*. Drug Safety (2020). <https://doi.org/10.1007/s40264-020-01028-w>
2. Gemzoe, K., Painter, J. L., Pimenta, J., & Powell, G. *Crowdsourcing as a novel method to assess the impact of drug exposure on Belimumab pregnancy registry enrollment*. In PHARMACOEPI-DEMOLOGY AND DRUG SAFETY (Vol. 28, pp. 436-436). Aug, 2019.
3. Jeffery L. Painter *A Day in the Life of a Data Scientist: Modeling the Impact of Subscriber Attrition with Variability*. <http://bit.ly/2rvbfri> Dec 13, 2017.
4. Schifano, L., Powell, M., Clayton, L., Akhtar, A., Painter, J., Jan, D., & Simard, E. P. *Discussions About Medication and Vaccine Use During Pregnancy and Lactation on Pregnancy Related Social Media Sites*. In PHARMACOEPIDEMOLOGY AND DRUG SAFETY (Vol. 26, pp. 511-512). Aug, 2017.
5. Alex Gartland, Jeffery L. Painter, Tim Casperson and Greg E. Powell *Leveraging Crowdsourcing to Help Classify Social Media Data for Medical and Patient Safety Insights*. AMIA Poster Presentation, Washington, DC, Nov 04-08, 2017.
6. Michele Thomas, Amy Curry, Jeffery L. Painter, Arhooj Akhtar, Lorrie Schifano, and Greg E. Powell *Case Study: Computing Complexity Scores to Identify Patients of Interest from Inspire.com Forums for Safety and Beyond*. DIA Poster Presentation, Chicago, IL, June 20, 2017.
7. Laurie S. Anderson, Heidi G Bell, Michael Gilbert, Christina Winter, Monica J Barratt, Beta Win, Jeffery L Painter, Christopher Menone, et. al. *Using Social Listening Data to Monitor Misuse and Nonmedical Use of Bupropion: A Content Analysis*. JMIR Public Health Surveillance, Feb. 2017.
8. Jeffery L. Painter *On the Perils and Pitfalls of PRR Analysis*. JSM, Chicago, IL, July, 2016.
9. Yingzi Xu, Jeffery L. Painter *Application of Classification and Clustering Methods on mVoC (Medical Voice of Customer) data for Scientific Engagement*. JSM, Chicago, IL, July, 2016.
10. Timothy A. Casperson, Jeffery L. Painter, and Juergen Dietrich *Strategies for distributed curation of social media data for safety and pharmacovigilance*. Proceedings of the International Conference on Data Mining, DMIN 2016, July 27, 2016, Las Vegas, Nevada, USA, CSREA Press
11. Gregory E. Powell, Harry A. Seifert, Tjark Reblin, Phil J. Burstein, James Blowers, J. Alan Menius, Jeffery L. Painter, Michele Thomas, Carrie E. Pierce et. al. *Social Media Listening for Routine Post-Marketing Safety Surveillance*. Drug Safety 2016. DOI: 10.1007/s40264-015-0385-6
12. Suzie Seabroke , Gianmario Candore, Kristina Juhlin, Naashika Quarcoo, Antoni Wisniewski, Ramin Arani, Jeffery Painter, Philip Tregunno, G. Niklas Norén, and Jim Slattery *Performance of Stratified and Subgrouped Disproportionality Analyses in Spontaneous Databases*. Drug Safety 2016. DOI: 10.1007/s40264-015-0388-3
13. Michele Thomas, MBA, Bradley Turnbull, PhD, Jeffery L Painter, BS, JD, Rachael L DiSantostefano, PhD and Greg Powell, PharmD, MBA. *What are People Talking About? Medication Discussions on Social Media*. ICPE, Boston, August 25, 2015, Poster session.
14. Greg Powell, James Blowers, Heidi Bell, Michele Thomas, Jeffery Painter, JD, Nabarun Dasgupta. *Comparison of Events in Spontaneous Adverse Event Reports To Events Discussed Within Context of Drug Use on Facebook and Twitter*. ICPE, Boston, August 25, 2015, Poster session.
15. Rachael L DiSantostefano, Jeffery L. Painter, Michele Thomas and Greg Powell *Safety Assessment and Selection Bias: Who uses social media to communicate about medications?*. ICPE, Boston, August 25, 2015, Poster session.

16. Matthew R Nelson, Hannah Tipney, Jeffery L Painter, Judong Shen, Paola Nicoletti, Yufeng Shen, Aris Floratos, Pak Chung Sham, Mulin Jun Li, Junwen Wang, Lon R Cardon, John C Whittaker and Philippe Sanseau. *The support of human genetic evidence for approved drug indications*. Nature Genetics, June 2015.
17. Daniel Parks, Xiwu Lin, Jeffery L. Painter, Jie Cheng, Christine M. Hunt, Colin F. Spraggs, Jeanenne J. Nelson, Lloyd Curtis, J. Alan Menius and Kwan R. Lee. *A proposed modification to Hy's law and Edish criteria in oncology clinical trials using aggregated historical data*. Pharmacoeconomics and Drug Safety, Volume 22, Issue 6, pages 571 - 578, June 2013
18. Jeffery L. Painter and Greg Powell. *Performance of Observational Screening by System Organ Class*. Pharmacoeconomics and Drug Safety, Vol. 21, ICPE 2012 Poster Presentation
19. Daniel Parks, Xiwu Lin, Jeffery L. Painter, Jie Cheng, Christine M. Hunt, HA Stirnaedl-Farrant, J. Alan Menius and Kwan R. Lee. *Validation of multivariate outlier detection analyses used to identify potential drug-induced liver injury in clinical trial populations*. Drug Safety, Volume 35, Issue 10, pages 865 - 875, Oct 2012
20. Cheng J, Greshock J, Painter J, Lin X, Lee K, Zheng S, Menius A. *Predicting breast cancer chemotherapeutic response using a novel tool for microarray data analysis*. Journal of Integrated Bioinformatics, Volume 9, Issue 2, pages 209, August 2012
21. Jeffery L. Painter. *An Imaging Framework for the Analysis of Longitudinal High-Dimensional Data*. International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2012, July 15-19, 2012, Las Vega, Nevada, USA, CSREA Press
22. Jeffery L. Painter. *Enhanced Biomedical Taxonomy Mapping Through Use of A Semantic Measure of Proximity*. International Conference on Knowledge Engineering, IKE 2011, July 2011, Las Vega, Nevada, USA, CSREA Press
23. Reisinger SJ, Ryan PB, O'Hara DJ, Powell GE, Painter JL, Pattishall EN, Morris JA. *Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases*. Journal of American Medical Informatics Association, Volume 17, Issue 6, pages 652-662, Nov-Dec 2010
24. Jeffery L. Painter. *Containing the Cloud: Security Issues in a Large Scale Observational Pharmacovigilance Research Project*. to be presented at the International Conference on Security and Management, SAM 2010, July 12-15, 2010, Las Vega, Nevada, USA, CSREA Press
25. Jeffery L. Painter. *Towards Automating an Inference Model on Unstructured Terminologies: OXMIS Case Study*. Advances in Computational Biology, Vol. 680, Ed. Arabnia, Hamid R., Springer-Verlag, 2010
26. Jeffery L. Painter and Natalie L. Flowers. *CodeSlinger: An Interactive Biomedical Ontology Browser*. Artificial Intelligence in Medicine: 12th Conference on Artificial Intelligence in Medicine in Europe, AIME 2009, Verona, Italy, July 18-22, 2009, Proceedings; Springer, 2009
27. G.H. Merrill, P.B. Ryan, and J. L. Painter. *Applying a UMLS/SNOMED-based Drug Ontology for Observational Pharmacovigilance*. IDAMAP (Intelligent Data Analysis for bioMedicine and Pharmacology), Washington, DC, November 7, 2008.
28. P.B. Ryan, G. H. Merrill, and J. L. Painter, *Defining medical conditions by mapping ICD-9 to MedDRA: A systematic approach to integrating disparate observational data sources for enabling enhanced pharmacovigilance analyses*, Drug Information Association, Boston, 2008, (Poster session.).
29. G. H. Merrill, P. B. Ryan, and J. L. Painter, *Using SNOMED to Normalize and Aggregate Drug References in the SafetyWorks Observational Pharmacovigilance Project*, KR-MED, Phoenix, AZ 2008, (Poster session.).



30. J. L. Painter, *A Mapping Between SNOMED-CT and the OXMIS Coding Scheme*, KR-MED, Phoenix, AZ 2008, (Poster session.).
31. J. L. Painter and G.H. Merrill. *Introduction to the UMLS for Text Analytics*, Technical Briefing for NC State Univeristy's Advanced Analytics Graduate Program. Presented September 12, 2007.
32. Kristoph Kleiner, Gary H. Merrill and Jeffery L. Painter. *Inter-translation of Biomedical Coding Schemes Using UMLS*. Proceedings of the 2006 AAAI Fall Symposium on Semantic Technologies.

## Interests

---

- Bridge
- Reading
- Patent Law
- Cooking