2441 Sheridan Avenue South
Minneapolis, MN 55405
+1 (612) 229-6779
brian.repko@learnthinkcode.com
http://www.learnthinkcode.com
https://www.linkedin.com/in/brianrepko

W Brian Repko

Bioinformatics software architect and technical leader with experience in biomedical research and pharmaceutical industry. Successful in leading projects to on-time and on-budget delivery. A positive influence on teams – focused on customer value, training and mentoring. Contributor to various open-source projects.

Professional Skills

Programming Languages	Java, Scala, Groovy, R, Python, Javascript, bash/csh, SQL, Ruby, Perl, C, C++
Java / JVM	Java SE, Java EE, Spring, Guava, Jersey, CXF, Hibernate, EclipeLink, Liquibase, Flyway
Bioinformatics / Data Science	tidyverse, shiny, htmlwidgets, BioConductor, BioJava, numpy, scipy, Jupyter notebooks, GATK, ADAM, GA4GH, various file formats and tools
Web and XML	HTML, CSS, JQuery, Angular, React, XML (various tech), REST and WS-*, RDF/OWL
Infrastructure Tools	*Hadoop* (Spark, Hive, Pig, HDFS), *Java EE servers* (Tomcat, Jetty, etc.), *Web servers* (Apache httpd, nginx, etc.), *Database servers* (Snowflake, Oracle, Vertica, Cassandra, MySQL, PostgreSQL, SQLLite, etc.), *Messaging* (ActiveMQ, akka, etc.), *Search* (Solr, ElasticSearch), *HPC* (GridEngine, modules, EasyBuild)
Development Tools	*IDEs* (IntelliJ IDEA, RStudio, Eclipse), *VCS* (Git, Subversion, CVS, Mercurial), *Build* (Maven, Gradle, Ant, Ivy), *Testing* (JUnit, TestNG, JMeter, Cucumber, FIT, Fitnesse, Selenium), Chef, *CI* (CruiseControl, Jenkins, Bamboo), JIRA, Bugzilla, XPlanner, various Wikis
Operating Systems	Apple macOS, Microsoft Windows, Linux (RHEL / CentOS, Ubuntu), Unix (various), VMS
Software Engineering	Agile, Lean, Scrum, XP, Crystal, Unified Process, UML, D-/T-/B-/AT-driven development

Professional Experience

Novartis Institutes for Biomedical Research (NIBR) – Oncology Engineering **10/2020 – present**

- Responsible for various R packages involved in RNA, cfDNA, scRNA, and Visium assay data QC and filtering as well as clinical trial data standardization and harmonization framework for establishing best practices of our analysts in re-usable code.

- Responsible for explaining data models for RWD (custom and standardized models) and potential bias in those datasets. Worked with datasets from two main vendors that include various -omic assay results.

Carrot Health (Health Informatics), Minneapolis, MN **– Sr Dir Data Engineering** **12/2018 – 10/2020**

- Working manager for team (7) responsible for the design and development of all data and feature engineering for a health informatics data science startup. Data included public and privately acquired consumer, market, geographic, and reference data as well as all customer data (membership, medical and Rx claims, quality measures, and surveys). Solution involved execution of custom and standardized SQL, R, python, and bash via Matillion on Snowflake platform – solving for customer variety and volume as well as data quality. Solution was completely hosted on AWS making use of various components (EC2, S3, SQS, SNS, Lambda, Batch). Eliminated most technical debt while both supporting a growing number of customers (400% revenue growth) and adding new modules and features to our products.

- Responsible for architecture and execution of all Carrot Health predictive models based on changing underlying data (DataOps). Responsible for testing and improvement of person-matching algorithms (record linkage).

- Wrote and contributed Snowflake support for both Flyway and SchemaSpy open-source tools.

Novartis Institutes for Biomedical Research (NIBR) 07/2011 – 10/2018

NIBR Informatics – Senior Principal Scientific Software Engineer / Genomics Lead
10/2015 – present, Minneapolis, MN *(remote to Cambridge, MA)*

NIBR Oncology Bioinformatics
02/2014 – 10/2015, Basel, CH

NIBR Developmental and Molecular Pathways (DMP) Computational Engineering
05/2013 – 01/2014, Basel, CH
07/2011 – 05/2013, Minneapolis, MN *(remote to Cambridge, MA)*

- Technical lead / agile coach for team (12) on data warehouse and analysis system holding all public and private –omic (genomic, epigenetic, transcriptomic, proteomic) experiment data. Apache Spark-based ETL framework with various outputs including Vertica database cluster. Support R package for data analysis and visualization as well as Jupyter notebooks.

- Responsible for various "*Engineering for Bioinformaticians*" initiatives including an AWS-based "Informatics Lab", Chef cookbooks for R, R libraries, Shiny Server, and Easybuild. Also oversee the rollout and training of key initiatives from other NX groups including migration to AWS/Turbot, DevOps tooling and multiple OS support (CentOS7, RHEL7, Ubuntu).

- Embedded engineer in the Oncology Bioinformatics team (7) in Basel – 1 of 4-person global Oncology engineering team. Oversaw transition of applications from Emeryville, CA team and integration of Translational Clinical Oncology in Basel. Served as liaison between Oncology and NIBR Informatics groups. Provided Oncology tool and application support and was main developer for the Oncology Global Data Share (R packages and data for all key experiments) and other key drug discovery and development project tooling. Provided engineering support for a knowledgebase of laboratory skills (based on JIVE).

- Senior software engineer on multi-tiered model organism genomic assembly and annotation reference system. System included a file system repository for all public (and private) genomic / proteomic data releases as well as post-processing of that data – standardized formatting, pre-built indices for various alignment tools (blast, bowtie2, tophat, star, wham, etc.). All internally used sequencing pipelines then used these files for the purposes of reproducible research. Data was also loaded into a customized GMOD Chado database. REST-based services (implemented in Java, Hibernate, Groovy, and Grails) were developed to support a "Gene Portal" biologist research tool.

LearnThinkCode (Consulting), Minneapolis MN – **President / Independent** 05/2006 – 05/2013

- Architect and agile coach on a customer-direct online health advocacy platform. Responsible for high-level designs and technical mentoring as well as rollout of agile methodology for web and mobile-based extension of core advocacy platform. Solution was Spring MVC-based web application with core services, based on the Spring Framework also exposed via REST (using CXF). Spring MVC-based web application rendered dynamic assessments using Metawidget framework and personalized coaching via FreeMarker.

- Agile coach and architect for multiple project teams within a multi-project program. Responsible for rollout of Agile methodology as well as high-level designs and technical mentoring for multi-year mission-critical rearchitecture program. Solution was a RIA with REST-based services utilizing Javascript, Flex and BlazeDS on the user interface and Jersey / Spring / JPA / Hibernate for the services running on Tomcat and Oracle.

- Lead technical architect for 60-developer organization. Agile coach and mentor for offshore development team. Lead for team (8) responsible for development of SaaS-based Content Management System. Senior Developer on a variety of initiatives – internationalizing a SaaS-based system; creating B2B Web Services using JAX-WS, WS-Security, CXF, Spring and Hibernate; and a Proof-of-Concept related to Google Translation. Training and mentoring of architects, developers and designers.

William Hill plc (Bookmaker), Leeds, West Yorkshire UK – **Agile Coach / Architect** **06/2007 – 01/2008**

- Agile coach and senior developer for three of five teams within a program (60 people) to implement William Hill's NextGen architecture. Promoted to Scrum Master and Technical Architect for the Online Sportsbook (bet slip) team. Solutions focused on full use of Spring Framework, Hibernate and Tangosol cache on WebLogic 9.2 and Oracle 10g. Testing tools included FIT, Fitnesse, Selenium and DBFit.

Education

Master of Arts
Mathematics
University of Wisconsin, Madison, WI

Bachelor of Science (Honors)
Computer Science, Mathematics and Philosophy
University of Wisconsin, Madison, WI

Certifications, Publications and Presentations

11/2017 **Product Agility When Your Product Is Curing Cancer** – Agile Days Twin Cities

07/2017 **Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening,** *Cell, 170(3), 577-592*

02/2016 **Engineering in the Disease Areas and Platforms** – NIBR Informatics, "Show-and-Tell"

01/2016 **Introduction to Bioinformatics for Engineers** – NIBR Informatics, 5-part training series

12/2014 **"Big Data" Bioinformatics** – NIBR Informatics, "Geek Speak"

01/2012 **Crucible** and **CFEngine 3** (2 talks) **-** CODE Freeze 2012, Minneapolis, MN

11/2011 **Introduction and Strategies for Effective ATDD -** Agile Day Twin Cities, Minneapolis, MN

01/2011 **Tests That Tell the Story of the Product** (session lead / part of panel)
CODE Freeze 2011, Minneapolis, MN

04/2010 **Test Driven Development (beyond JUnit) with JBehave**
Java Users Group, Minneapolis, MN; Milwaukee, WI; and Madison WI

06/2008 **HTTP Push Technologies** - Java Users Group, Minneapolis, MN

06/2005 **Spring Application Framework - Java Users Group, Minneapolis, MN**

04/2005 **Application Performance with JAMon and JMeter - Java Users Group, Minneapolis, MN**

10/2003 **Web Services - Separating the Hype from Reality - CIO Roundtable, Cleveland, OH**

04/2003 **Improving the Harley-Davidson Supply Chain**
DCI / Meta Web Services Conference, San Francisco, CA