

Sashi Kiran Challa

720 S College Mall Road, F4 Hunter Ridge Apts, Bloomington, IN-47401

schalla@indiana.edu | (812) 606-3254

PROFILE

- Ability to program in Java, Perl, Python, Shell-Scripting, C#, R, SAS.
- Familiar with Software Development Life Cycle and Software Testing concepts, Next-Gen sequencing technologies, Machine Learning Techniques like Hierarchical Clustering, Probabilistic Modeling, Hidden Markov Models, Multiple Linear Regression, Genetic Algorithm, and Principal Component Analysis.
- Passion to learn and implement new technologies. Experience in using Semantic Web technologies and High Performance Computing technologies in the areas of Cheminformatics and Bioinformatics. And also experience in using Dryad (DryadLINQ); Microsoft's cloud runtime.
- Highly motivated, positive individual with Strong leadership, team building qualities and excellent briefing and presentation skills.

EDUCATION

Master of Science in Cheminformatics

Indiana University, Bloomington, IN

GPA: 3.78/4.00

May 2010

Master of Science in Honors Chemistry

Sri Sathya Sai University, Andhra Pradesh, India

GPA: 4.59/5.00

March 2007

Bachelor of Science in Honors Chemistry

Sri Sathya Sai University, Andhra Pradesh, India

GPA: 4.69/5.00

March 2005

EXPERIENCE

Center for Genomics and Bioinformatics,

Indiana University, Bloomington, IN

July 2010 - Present

Bioinformatics Programmer/Analyst

- Next-Gen Transcriptome data analysis; Comparative Genomics Project (3 different singing birds).
- Write Perl Scripts to do quality check on the reads obtained from Roche 454 sequencing machine.

Community Grids Lab, Pervasive Technology Institute,

Indiana University, Bloomington, IN

August 2009 - Present

Cheminformatics Programmer/Analyst

- Built a pipeline as part of the Microsoft Research's ORECHEM project, to process data and do computations.
- Worked with Dryad, DryadLINQ to distribute and run parallel jobs on different nodes in a windows cluster.
- Developed individual RESTful web services (using JAX-RS Jersey API) that (i) fetch ATOM feeds (ii) RDFize ATOM feeds (GRDDL, Saxon-he) and push them into triple store (WEBDAV client) (iii) extract relevant information (JDOM, JAXP, ROME API) from the ATOM feeds, perform relevant transformations and submit that data to structure modeling tool running on HPC clusters (TeraGrid). Built Test cases in JUnit for testing these RESTful web services.
- Successfully set up Open Link Virtuoso Triple Store on Linux and Windows platforms.
- Worked with XBay workflow composer to compose a workflow out of these REST services.

Persistent Indiana Research Center, Bloomington, IN

January 2010 – May 2010

Bio/Cheminformatics Programmer

Worked with a Client of Persistent Systems on incorporating Hierarchical Clustering, PCA components into a cloud based RT-PCR data analysis tool.

- Implemented Hierarchical Clustering algorithm in Java.

- Successfully was able to use R-Serve to make calls to R from Java, perform Hierarchical clustering, PCA in R, and get Cluster Heatmaps and Biplots (PC1vsPC2vsPC3 plots) into Java. Also worked with Bioconductor package in R.

**The Rieseberg Lab, Plant Genomics Lab,
Indiana University, Bloomington, IN
Web Programmer**

June 2009 - August 2009

- Worked with Plant Genomics group from University of British Columbia, Canada and contributed to development of EvoPipes Bioinformatics pipelines suite (<http://evopipes.net/>). Specifically wrote PERL/CGI scripts that take in input files from user, check for correct file format (fasta format) store it on the server, and run a bioinformatics pipeline.

RELEVANT CLASS PROJECTS

Genetic Algorithm for Feature Selection:

- Wrote a function in R that uses MLR to build models and performs Genetic Algorithm to select best descriptors for a QSAR model and examined how fast the genetic algorithm reaches the global best model. R^2 and RMSE were the objective functions used.

Functional Genomics and Metagenomics:

- Contributed in studying the cause for differences in early developmental patterns in two Sea Urchin species belonging to *Heliocidarus* genus; *H.erythrogramma* and *H.tuberculata*.
- Analyzed real time genome data provided by Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN. Got to use several bioinformatics tools like NCBI ORF Finder, Glimmer, GeneMark, AmiGO, PFP, ClustalW, HMMER, Phylip.
- Implemented gene finding algorithms that use Probabilistic modeling, 1st order Markov models (in Python). Learnt Phylogenetic Analysis algorithms.
- Contributed in comparing species diversity estimation in 13 human gut-microbiome metagenomes using 16SrRNA and recA phylogenetic markers.

Rules on Drug Discovery Data using Semantic Web Technologies:

- Built OWL Ontology for WENDI (Web Engine for Non obvious Drug Information), a tool developed for Eli Lilly Pharmaceutical Company by Cheminformatics group at Indiana University.
- Automated generation of RDF triples from XML output (using JAXP, JDOM in Java), based on the Ontology.
- Framed rules (using Jena's generic reasoner) to infer relationships between compounds and genes and diseases.

CONFERENCE ACCEPTANCES AND PUBLICATIONS

- Challa, S. Wild, D, J. Ding, Y. Zhu, Qian. Semantic Rules on Drug Discovery Data. 4th Annual Asian Semantic Web Conference (ASWC) 2009. Accepted October 2009
- Zhu, Qian. Challa, S. Purohit, P. Lajiness, M. Wild, D.J. and Ding, Y. "Using Web technologies for integrative drug discovery". Accepted at IEEE/WIC/ACM WI 2010
- Michael S. Barker, Katrina M. Dlugosch, Louie Dinh¹, R. Sashikiran Challa, Nolan C. Kane¹, Matthew G. King, and Loren H. Rieseberg. "EvoPipes.net: Bioinformatics tools for ecological and evolutionary genomics". Evolutionary Bioinformatics. Manuscript Accepted Aug 2010.

TECHNICAL SKILLS

Languages: Extensive experience in Java, Python, Perl, Shell Scripting. Familiar with C, C++, C#

Databases: MYSQL, Postgre SQL, MS Access

Web Technologies: PHP/MYSQL, PERL/CGI, HTML, CSS, XML, XSLT

Semantic Web Technologies: RDF, SPARQL, OWL (Ontology Engineering), D2RQ mapping.

Frameworks: Jena, SILK-LSL, WAMP, GATE (Text Engineering)

Web-Apps Hosting: Apache Tomcat, Google App Engine

Statistical Packages: Extensive experience in R. Familiar with BASE-SAS, SAS-GRAPH, SAS-STAT

Platforms: Comfortable with Open Eye toolkit, Pipeline Pilot. Also Comfortable on Various OS Platforms (RedHat/CentOS/Ubuntu Linux, OS X, Windows)

Version Control Systems: Familiar with SVN, Mercurial

IDE/Simulators/Build tools: Eclipse, Microsoft Visual Studio 2008, Protégé, Adobe Dreamweaver, Maven, Ant