**Profile**

* Recent Ph.D. graduate looking for a career in Data Science or Software Engineering

**Education**

**B.Sc. Computer Science** (Minor Statistics)University of British Columbia *2006–2011*

**Ph.D. Bioinformatics**University of British Columbia *2011–2015*

**Coursework**

* Machine Learning, Data Visualization, Numerical Methods, Algorithms, Software Development

**Technical Skills**

* Programming Languages Java, C++, R, Python, JavaScript, and PHP
* Object-oriented software design
* Cloud data processing experience using AWS and Hadoop
* HPC programming using Intel Phi-cards
* Data Analysis using the ggplot2, RMarkdown, and d3.js

**Experience**

**Bioinformatics Ph.D. Candidate** *2011-2015**Hallam Lab, University of British Columbia*

* Developed MetaPathways, a Python and   
  C++ pipeline for next-generation sequencing data
* Implemented a Master-worker distributed algorithm for managing a collection of non-cooperative HPC worker grids.
* Processed next-generation sequencing information in Hadoop via improvements to the YARN scheduler
* Involved in analysis of all research; author on 10+ peer-reviewed publications

**Teaching Assistant**  *2014**Marine Biological Laboratory, Woods Hole, MA*

* Conducted tutorials on sequence analysis for 60+ post-docs, graduate students, and professors
* Statistical topics: Linear Models, SVMs, Hierarchical Clustering, PCA, NMDS
* Data Visualization in R: lattice, ggplot2, and RMarkdown

**Software Developer** *2011*  
*Microbiology and Immunology, UBC*

* Developed a database-driven web-application for bacterial strains
* Utilized the LAMP stack, JavaScript, and PHP to create an easy-to-use web interface
* Implemented usability improvements via an advanced search form, user search histories, and intelligent autocomplete

**Projects**

* **De-Confusion Tables**. An interactive R shiny app. Final project. Developing Data Products. John Hopkins University Coursera Class. July 2014 (<http://goo.gl/Fl78Qh>).
* **D3.js Visualizations**. Variety of visualizations in the d3.framework for analysis of genomic data (<http://goo.gl/fLntUc>).
* **R-Package: fastLCA**. Implementation of correlation method fastLCA in C and Open-MP (<http://goo.gl/5QPM82>).

**Selected Publications**

For references see Google Scholar (<http://goo.gl/aYo3Wa>).

* Kim, Konwar, **Hanson**, *et al.* Koonkie: An Automated Software Tool for Processing Environmental Sequence Information using Hadoop. ASE BigData. Harvard. (2014)
* **Hanson**, *et al.* MetaPathways v2.0: A master-worker model for environmental Pathway/Genome Database construction on grids and clouds. MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. IEEE CIBCB, Hawaii (2014)
* Konwar, **Hanson**, *et al.* MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. BMC Bioinformatics (2013)

**Journal & Conference Reviewing**

* IEEE Conferences CIBCB 2014 and 2015
* Symposium on Bioinformatics Research and Applications ISBRA 2015

**Interests**

Data Analysis, Information Visualization, Machine Learning, Statistics, High-performance Computing, Web-app Creation, Software Design

**References**

Provided upon request.