

Dr. Victoria A. Stuart, Ph.D.
Cover Letter

Vancouver, British Columbia, Canada
(604) 628-9609
mail@VictoriasJourney.com
Persagen.com

Background

As indicated in my Curriculum vitae (<https://persagen.com/about/victoria/cv.pdf>), I have a broad wealth of knowledge, expertise and experience.

**Undergraduate
Education**

My undergraduate degree in Biochemistry introduced me to the fundamentals of biochemistry and metabolism. I also worked as in the laboratory of Dr. Robert W. Chambers, a Nobel Prize candidate and the first recipient of the Terry Fox Special Initiatives in Cancer, through which I acquired high-level expertise in organic chemistry (polynucleotide synthesis, culminating in a Nucleic Acids Research publication), cancer biology, and microbiology.

**Graduate
Education**

For my graduate studies I first completed a M.Sc. in Occupational Hygiene (Environmental Health), focusing on the genotoxic effects of an agricultural pesticide. As part of that work I wrote a script to process florescent microscope images, that I directly imaged onto the CCD chip of the digital camera.

I followed those studies with my doctoral studies, under the supervision of noted biologist / molecular geneticist Dr. Barry Glickman, in which I investigated the mechanisms of mutagenesis and mutagenesis. In those studies I characterized the origins of spontaneous mutations with age (associated, e.g., with cancers and other age-related illness), as well as the tissue specificity of dietary mutagens and carcinogens. During this period I also directed technical staff and summer students who assisted my work, and taught some classes relating to DNA repair — coordinating one course. This work, which I presented at many conferences, resulted in a dozen or so publications in leading journals.

**Postdoctoral
Work**

I continued those molecular genetics studies during my postdoctoral work at the National Institute of Environmental Health Sciences in the laboratory of the scientist who cloned the human mitochondrial DNA polymerase (Dr. Bill Copeland) – cosupervised by Dr. Micheline K. Strand. My work focused on mitochondrial mutations involved in DNA repair and human mitochondrial-based diseases, using yeast as a model organism.

During this time I was awarded a prestigious National Academy of Sciences Research Associate Program award, that extended my stay allowing me to investigate genomic changes in expression in the yeast genome as a consequence of mtDNA mutations and exogenous stressors (e.g. oxidants).

In one of my studies I noted a curious pleiotropic effect (described more fully in my Omnigenics research blog post: <https://persagen.com/2020/02/05/omnigenics.html#overlooked>), highly complex interactions among yeast transcription factors pleiotropic effects on global cellular regulation. Although my work was not cited, in 2017 Dr. Jonathon Pritchard and colleagues at Stanford University derived similar conclusions regarding the genomics underlying complex phenotypes.

**Personal
Projects**

Since my undergraduate days I have held a profound interest in all aspects of scientific exploration and discovery – particularly with our genetic code, DNA. That interest was sparked, in part, by Watson and Crick's elegant one-page paper (Nature, 1953) describing the DNA helix. While conducting my undergraduate studies, I surveyed all of the books in my Supervisor's

office and the departmental reading room (as well, e.g., the autobiographies published by leading scientists in the various Annual Reviews of * series), providing a deep appreciation of the scientific method generating that knowledge.

Likewise, my postdoctoral work sparked a deep fascination with bioinformatics: the discovery, processing and interpretation of genetic-encoded information. I have a long-held desire to collect, collate and understand biomedical knowledge. To address those aims, Over the course of several years I acquired the programmatic skills described elsewhere in this cover letter and my curriculum vitae, while continually refining my approach. My most recent work focuses on processing text, extracting relationships, and analyzing and visualizing those data.

Addressing Disinformation

We live in an era where disinformation – which has monetary and persuasive value – proliferates throughout society including our scientific publications. The consequences of disinformation are dire. For example, anti-vaccination disinformation results in loss of "herd immunity," as seen recently in Samoa. Critics of climate change result an existential threat to the millions (perhaps billions) of people directly and indirectly affected by climate change. Predatory publications in biomedical and toxicological journals likewise affect medical understanding, interpretation and practice, affecting lives – e.g.: interference with the regulation of pesticide safety; recommendations regarding the dietary consumption of fat | sugar; "dangers" of vaccination; disinformation on the dangers of opioid pharmaceuticals; ...

Society (general; scientific; technical) requires factual, accurate information to drive the decisions that are needed to address challenges facing societies, including climate change, ecological destruction, health decisions and policies, wealth inequality and disproportionate influence and power, and emergent challenges including automation and other economic challenges posed e.g. by machine learning and artificial intelligence.

The foundational knowledge needed for rational decision-making is increasingly eroded in a disinformation, post-truth era – with notable examples being the rescinding of environmental protections and increased hate and xenophobia in the USA, driven in part by disinformation sources such as Breitbart and other entities.

Proof of Concept

As proof-of-concept I am currently investigating the sources and effects of disinformation, constructing a knowledge graph that is free (as much as possible) from that disinformation. The current domain for that work is the general discourse presented e.g. in the new media. In addition to developing tools and approaches to address disinformation, another reason for this work is that natural language processing tools are semantically well-suited to this domain (aiding named entity recognition, disambiguation and entity linking, for example).

A strength of the approaches and tools that I am developing is that they are broadly applicable to many domains. For example, part of my work includes creating ontologies, that enable accurate information extraction in specific domains, including biomedical research.

Personal Statement / Vision

I seek innovative, impactful solutions to scientific and societal issues through the application of science and technology to the betterment of human health. These include the following areas.

(i) The advancement of knowledge via information extraction in core domains (biomedical; social; environmental; ...) via natural language processing and machine learning approaches.

(ii) The discovery of factual and accurate relationships – explicit and implicit – among those data through the construction and use of graphical models, and data visualization.

(iii) Validating sources of information, fact-checking of existing relationships, and identify-

ing and removing sources of disinformation.

These efforts build on my thorough grounding in biochemistry (B.Sc.), environmental health (M.Sc.), molecular genetics (Ph.D), and post-doctoral experience in bioinformatics.

Recent Work

Scientific Consultant | Owner: Persagen.com

Jun 2009 – Present

Subcontractor

I have been self-employed since June 2009 as a Scientific Consultant, providing scientific expertise in molecular genetics, genomics, molecular biology, life sciences, bioinformatics and scientific review.

From 2009-2014 I was subcontracted to Battelle Memorial Institute (Chapel Hill, N.C.), providing scientific expertise for the U.S. Army Research Office (ARO; Durham, N.C.) and the U.S. Army Center for Environmental Health Research (USACEHR; Washington, D.C.). That work included evaluation of molecular genetics and other life sciences research proposals for the ARO, and analyses of genomic expression data (*Yersinia pestis*-infected primate cells) for the USACEHR.

That work, with my previous postdoctoral studies, reinforced my desire to acquire expertise in (bio)informatics.

Self-Supported Studies

I am fundamentally interested in information and the relationships among data. To accomplish the aims expressed in my Personal Statement and elsewhere in this cover letter, from 2015 – present I have been immersed in a highly concentrated self-supported and -directed programme of study to acquire the tools in core domains (data management; programming languages; informatics; graphical models; data visualization; natural language processing; machine learning) needed to realize my Vision.

I possess a theoretical understanding and appreciation of the underlying principles (modeling; loss functions; gradient descent; back propagation; tensors; linear algebraic approaches such as eigenvalues and semirings applied to data matrices; etc.) associated with leading-edge machine learning and natural language processing models.

Although I have not yet deployed leading edge contextual language models and machine learning approaches (e.g. representation learning and link prediction; hyperbolic embeddings) in my current work, I follow these domains with the intention of investigating their use in the next stages of my research.

I am presently concentrating on the use of established, yet high performing approaches to relation extraction (e.g. shortest paths dependency parsing) and graph construction (standard and bespoke ontologies; semantic entity linking; RDF triples; ...), for characterizing textual data sourced from the web, the biomedical literature, as well as a custom corpus of biomedical-related content spanning several hundred topic areas.

Programmatic Contributions

I have described some of my work in various research blog posts on my website, Persagen.com.

Please also note my contributions from 2013 – present to the programming community, e.g.

- StackOverflow: user:1904943 | ~2k reputation | top 7%, 2019
- comments and solutions to GitHub Issues; etc.