

Gene Selkov

86 Wellbrook Way
Girton, Cambridgeshire CB3 0GP

Phone: 07501461257

Email: selkovjr@gmail.com

Email: gs437@cam.ac.uk

Employment History**Programmer**

Zoology, University of Cambridge

February 2011–Present

Develop a web front end to ACMACS (Antigenic Cartography Macros), a novel system for visualizing the relationships between antigenically variable pathogens.

This job involves using 15-year-old technologies now finally available in the web browsers to satisfy the demands of a futuristic research environment. It is a work in progress.

Programmer

Self-employed

June 2010–Present

Collaborate with CDR Paul Roach of the U.S. Navy on the next-generation **STAT** (Surgical Training and Assessment Tool). Maintain the running installations of **STAT** at the University of Chicago Hospitals, Saint Agnes Hospital, Naval Medical Center Portsmouth, and a number of other institutions.

Collaborate with Dr. Sam Volchenbom of the University of Chicago Center for Advanced Medicine on the development of new algorithms and experimental designs for real-time proteomics.

The goal of the latter work is to enable fast and non-redundant data acquisition from a mass spectrometer to allow reliable identification of human proteins in real time. Because the software supplied with the instrument is inadequate for the level of control this work requires, and the instrument vendor refuses to disclose its source code and data specifications, the only way to improve it is by reverse-engineering the instrument and its data encoding. The resulting knowledge of data layout and practical decoder implementations for several versions of Thermo-Finnigan data format are being released to the public through Google Code.

Senior Software Developer

Computational Systems Biology, The University of Edinburgh

June 2007–May 2010

In collaboration with the University of Chicago, completed **STAT**, a web-based training and assessment system for surgical fellows and residents. Designed and built a web-based area survey system for the South Side Health and Vitality Studies in Chicago. Conducted a feasibility study of the application of Mushkambarov monocarbon fragment analysis to the energy assessment of whole-genome metabolic reconstructions. Participated in the development of **ts**, a *PostgreSQL* extended datatype to enable massive storage and similarity-based retrieval of time series realizations.

Research Systems Analyst

Computation Institute, The University of Chicago

November 2004–June 2007

Designed a prototype for a web-based system for progress tracking and competency assessment of in the training of surgeons (**STAT**); performed statistical analyses of the training program's performance. Developed algorithms and built a web-based toolchain to support the genotyping of microbial communities. Developed a method for visualizing trajectories of palpation in simulated pelvic exams. Developed tools for remote collaboration based on Access Grid (venue recorder, venue topology mapper and telemetry applications, including physiological monitors). Built a fast GPU-based volume and surface rendering cluster for the Visible Human Project by the National Library of Medicine. Developed a process for anaglyph rendering of 3D images, movies, and interactive visualizations. Designed and built visualization hardware. Designed a database to support the Lung Cancer Tissue Bank. Helped Dr. Yasushi Nakagawa establish the genetic identity of the kidney stone inhibitor he discovered.

Senior Scientific Programmer

Argonne National Laboratory

August 2002–Present

Set up and outfitted a new microbiology lab. Developed procedures and built equipment for storage, cultivation and synchronous growth of photosynthetic bacteria in automated bioreactors. Provided technical support and conducted experiments in search of evidence to support Evgeni Selkov Sr.'s theory of circadian oscillations. Developed web-based

monitoring and control interfaces for bioreactors, enabling remote collaboration of the participating researchers.

Systems Designer

Integrated Genomics, Inc.

January 1999–March 2003

Designed the network and computational infrastructure for the company during its startup period. Procured contracts with internet service providers. Set up and maintained a SKIP-based virtual private network for the telecommuting employees. Carried out UNIX installations and provided systems support. Participated in the development of the core product, **ERGO** -- an integrated system for comparative genome research and metabolic reconstruction. Developed **ERGO-Lite** -- a set of query tools to aid in comparative studies of multiple related genomes. Delivered products and provided training and support to customers. Designed a multiple-dispatch framework for serializable compute jobs on a Linux-based cluster. Created and maintained a comprehensive database of chemical names and structures of metabolites.

Visiting Scientist

Argonne National Laboratory

June 1998–June 2000

Designed a novel data type and access methods to represent the results of measurements as variable-precision numeric intervals suitable for storage in relational databases, and implemented it as a **PostgreSQL** extension, **seg**. Used **seg**, as well as other original data types, in the first relational indexing of the enzymology database, **EMP**. Developed query interfaces to **EMP** and to the metabolic pathway database, **MPW**. Developed a method, a markup language, and a prototype software tool, **ElectricArc**, to create and maintain diagram-based databases, such as metabolic pathways, electrical circuits, or any type of data that can be represented as a graph. **ElectricArc** served as a prototype for **Edinburgh Pathway Editor** now being developed at the University of Edinburgh.

Software Engineer

Laboratory of Mathematical Modeling, Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences

March 1993–May 1998

Provided systems and hardware support for the nascent database on enzymes and metabolic pathways, **EMP**. Developed recognition algorithms and data models to capture the content and connectivity of pseudo-graphical pathway diagrams. Developed and published the tools for parsing multi-level multiple lookahead grammars based on cascaded LALR(1) parsers; among them, a version of **GNU Bison** adapted to generate Perl code and Perl extensions to encapsulate **Flex**-generated scanners. Automated the transformation of the database of enzyme reaction mechanisms, **ERM**, originally encoded in Cleland notation, into algebraic-differential equations.

Research Consultant

Michigan State University

June 1997–August 1997

Collaborated with the **MSU Center for Microbial Ecology** on the integration of dozens of diverse microbial databases into a common search and retrieval system with a single interface and standardized query capabilities. Created adaptive parsers for each participating database to translate their flat-text content to a structured representation suitable for relational indexing. It was the first time that such well known data sources as **ATCC** and **Bergey Manual** could be queried online. The Ribosomal Data Project (**RDP**) used the parser I created for them to completely convert their data to a structured format they have since adopted as standard.

Vice President/Engineering

Resonance LLC, Pushchino, Russia

October 1990–March 1993

Managed a group of engineers developing high-performance encephalographic and cardiographic equipment. Developed data acquisition and brain mapping software for encephalography. Supervised manufacturing and procurement. During the period of hyperinflation in Russia (1991-1992), supported R&D and provided cheap components for the manufacturing by setting up a recycling operation that salvaged precious metals and reusable parts from old mainframe computers. Constructed the tools for rapid disassembly of PCBs and set up a number of high-yield metallurgical and electrochemical processes for the separation of metal alloys and for the purification of gold, palladium, silver, copper, lead and tin.

Junior Researcher (part-time)

Center for Advanced Nonlinear Modeling, Moscow, Russia

August 1989–December 1991

Assisted in the development and prototyping of distributed finite automata for rapid image recognition based on autowave computing. The systems built included target discriminators for missile guidance and in-transit scanners of railroad car markings. Created the TTL-based prototypes of compute elements for such systems.

Engineer

*Laboratory of Acoustics, Institute of Biophysics, Russian Academy of Sciences
March 1986–August 1989*

Developed computational methods for the rapid assay of the physiological status of animal tissues based on their macromechanical properties, such as compressibility and viscosity. Prototyped a microprocessor-based ultrasonic device for the evaluation of protein and fat content in milk. Prototyped a precision 3-D thermostat with a distributed array of heating elements controlled by an array of sensors. Created numerical models of acoustic interferometers

Computer Engineer (part time)

*Laboratory of Neurophysiology, Institute of Biophysics, Russian Academy of Sciences
August 1983–March 1986*

Installed and configured **IN-1200**, a French-made signal processor that was delivered with a rudimentary disk operating system, a BASIC interpreter and a primitive FORTRAN compiler, accompanied by voluminous but poor documentation. Translated and improved the documentation, trained the physiologists in the basics of programming and signal processing, and assisted in experimental design. Earned a Master's degree for porting a C compiler and a version of BSD UNIX to this machine via intermediate byte code.

Laboratory Technician

*Laboratory of Acoustics, Institute of Biophysics, Russian Academy of Sciences
July 1978–August 1980*

Provided technical support and tools for a series of biological experiments. Operated the laboratory machine shop. Constructed a high-intensity non-resonant ultrasonic disintegrator for tissues and fluids with a frequency-sweeping capability.

Miscellany

I have been involved in many other activities -- both for-profit and voluntary ones.

As a student, I was contracted by a naval research center to screen and review the foreign literature on deep sea diving. The information I found helped the divers to improve their body temperature control, reduce the overheating of the brain, and eventually surpass the 600 meter depth record held by the French Navy for decades (the military are not unlike sportsmen; they engage in their own unpublicised Olympics).

Because of my above-average language skills and familiarity with many areas of engineering, I was contracted by several institutions to translate technical documentation from English, German and French, and helped scientists publish their research articles in English.

While still a student, I was known as a demolition expert. I developed and used an original method of non-explosive dismantling of failing smokestacks in confined environments, using inflated tire tubes as internal scaffolding and pressure source.

For several years after graduation, I taught Physics at a high school as a volunteer. I also worked as a private tutor in Physics and English.

I ran a private computer repair business during the late 1980s, when personal computers and their components were too expensive to replace.

I am a free software activist. I respond to questions in the following fora:

pgsq-general@postgres.org

pgsq-sql@postgres.org

comp.lang.perl.tk

I have contributed to *PostgreSQL Programmer's Guide* and *PostgreSQL Administrator's Guide*

I am the author of *Unfinnigan*, an open-source decoder for proprietary mass-spec data formats.

Education

M.Sc. in Applied Mathematics and Computer Science, March 1986

Obninsk Institute for Nuclear Power Engineering, Russia (formerly a branch of Moscow Engineering and Physics Institute)

Computer Experience

Programming Languages: Perl, C, JavaScript, sh, tcl, SGML-derived languages(XML, SVG, HTML, publishing with Docbook).

Development Techniques: Object-Oriented Design, Literate Programming, GNU build system, generation of automata in various programming languages with Lex and Yacc (Flex and Bison), image processing, digital photography.

Tools: emacs, vi, Apache, nginx, node.js, PostgreSQL, Octave (Matlab), R, GIMP, Cinpaint, Inkscape, Blender.

Other Software: TCP/IP networking, UNIX programming.

Operating Systems: UNIX varieties: Linux, FreeBSD, OpenBSD, Solaris, OSX.

Hardware: Intel x86 architectures and peripherals, Routers, Microprocessor programming and emulation, Lab interfaces: CAMAC, VMBus, GPIB, UEI PowerDAQ.

Databases

data design and normalization, text processing, pattern recognition and filtering, logical programming, database administration, query optimization.

Web Development

Asynchronous interactive user interfaces based on the YUI toolkit and SVG graphics.

Engineering and Manufacturing Experience

Engineering Graphics: Drafting, CAD/CAM.

Electronics: Analog/RF circuit design, Digital design: microprocessors, computer peripherals, robotics, Switched-capacitor filtering, PCB layout and routing.

Lab work and machine shop work: casting, forging and welding, plastic molding, milling, turning, grinding and shaping, sheet metal forming, electro-etching and electroplating, glass blowing, vacuum technologies.

Publications

- Anatoly Sorokin, Gene Selkov, Igor Goryanin. "A user-defined datatype for the storage of time series data allowing efficient similarity screening". *European Journal of Pharmaceutical Sciences*. July 2012. 46(4): 272-274. <http://www.sciencedirect.com/science/article/pii/S0928098711004404>.
- Gene Selkov. "Unfinnigan: specifications and open-source tools for decoding Finnigan data files". *Google Code*. 2010. <http://code.google.com/p/unfinnigan/>.
- Paul B. Roach, Kevin K. Roggin, Eugene Selkov Jr., Mitchell C. Posner and Jonathan C. Silverstein. "Use of a Novel, Web-Based Educational Platform Facilitates Intraoperative Training in a Surgical Oncology Fellowship Program". *Annals of Surgical Oncology*. May 2009. 16(5): 1100-1107. <http://goo.gl/1cqXU>.
- Jonathan Silverstein, Gene Selkov, Lawrence Salud and Carla Pugh. "Developing Performance Criteria for the e-Pelvis Simulator Using Visual Analysis". *Studies in health technology and informatics*. February 2007. 125():436-438. <http://goo.gl/yw3cB>.
- Selkov, E., Selkov, G., Karpeev, D., DeHorn, D., Minkoff, M., Nikolaev, E., Poluektova, Y. "Metabolic theory of the cell clock and its experimental verification in sequenced cyanobacteria". *Abstr #822, In: Plant Biology 2003 Final Program, Honolulu, Hawaii, ASPB*. <http://abstracts.aspb.org/pb2003/public/P50/0704.html>.
- Overbeek R, Larsen N, Walunas T, D'Souza M, Pusch G, Selkov E Jr, Liolios K, Joukov V, Kaznadzey D, Anderson I, Bhattacharyya A, Burd H, Gardner W, Hanke P, Kapatral V, Mikhailova N, Vasieva O, Osterman A, Vonstein V, Fonstein M, Ivanova N, Kyrpides N. "The ERGO genome analysis and discovery system". *Nucleic Acids Research*. January 2003. 31(1): 164-71.
- A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, Bernal A, D'Souza M, Ivanova N, Kapatral V, Larsen N, Los T, Lykidis A, Selkov E Jr, Walunas TL, Purcell A, Edwards RA, Hawkins T, Haselkorn R, Overbeek R, Kyrpides NC, Predki PF. "Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights". *Genome Research*. October 2002. 12(10): 1556-63.
- Bruce Momjian, Gene Selkov, Jr. "PostgreSQL: Advanced Topics Tutorial". O'Reilly Open Source Convention, San Diego, CA. July 2001.
- R. Overbeek, N. Larsen, G.D. Pusch, M. D'Souza, E. Selkov, N. Kyrpides, M. Fonstein, N. Maltsev, and E. Selkov, Jr. "WIT: integrated system for high-throughput genome sequence analysis and metabolic reconstruction". *Nucleic Acids Research*. January 2000. 28(1): 123-125.
- Selkov, E. Jr., Gretchkin, Y., Mikhailova, N., Selkov, E. "MPW: The Metabolic Pathway Database". *Nucleic Acids Research*. 1998. 26(1), 43-45.

- Selkov, E., Galimova, M., Goryanin, I., Gretchkin, Y., Ivanova, N., Komarov, Y., Maltsev, N., Mikhailova, N., Nenashev, V., Overbeek, R., Panyushkina, E., Pronevitch, L., Selkov, E. Jr. "The Metabolic Pathway Collection: an Update". *Nucleic Acids Research*. 1997. 25(1), 37-38.
- Selkov, E., Basmanova, S., Gaasterland, T., Goryanin, I., Gretchkin, Y., Maltsev, N., Nenashev, V., Overbeek, R., Panyushkina, E., Pronevitch, L., Selkov, E. Jr., Yunus, I. "The metabolic pathway collection from EMP: the Enzymes and Metabolic Pathways Database". *Nucleic Acids Research*. 1996. 24(1), 26-28.
- Sarvazyan, A. P., Selkov, E. E. Jr., Tchalikyan, T. B. "A constant length acoustic interferometer with a transition layer for precision measurements in small volumes of liquids". *Akusticheskii Zhurnal (in Russian)*. 1988. 34(6), 1103-1106.