3600 Market Street Suite 810
Philadelphia, PA 19104

Stephen M. Grimes, Ph.D.
http://www.ldc.upenn.edu/~sgrimes

sgrimes@ldc.upenn.edu
(812) 650-4466

_____

SOFTWARE, DATA, AND COMP LING EXPERIENCE

Linguistic Data Consortium, Philadelphia, PA   2012-present
Senior Software Developer, team leader
- Supervise software development team focused on backend data management for DARPA sponsored projects including MADCAT, DEFT, KBP, BOLT, HAVIC, GALE, and RATS.
- Designed XML data formats and conducted expert quality control of information content
- Standardized data delivery patterns at LDC using rakefiles and git code libraries
- Implemented workload balancing, knowledge redundancy, and team-driven development using collaboration methods such as pair programming, code review, and agile development philosophy.
- Recruited, vetted, hired, trained, and software developers
- Contribute to Rails webapp by customizing views using jquery and javascript

Linguistic Data Consortium, Philadelphia, PA   2008- 2012
Programmer Analyst / Application Developer
Word alignment project lead
- Designed and engineered desktop annotation GUI using Python and Qt library to record gold standard word alignments between parallel sentences
- Leveraged Posterior Regularization machine learning methods to automatically label word alignment relations in an effort to speed manual annotation
- Designed innovative Parallel Word Aligned Treebank corpus for machine translation research
MADCAT handwriting annotation project tech lead
- Managed Codeigniter/PHP/JavaScript webapp used to annotate handwriting collections for use as OCR training
- Created data integrity tests find inconsistencies in subcontractor data

American Indian Studies Research Institute, Indiana University   2005-2008
- Enhanced Dictionary Database and Automated Text Processor software programs used in creating and managing linguistic resources of American Indian languages
- Created fonts and converted language data between fonts encodings
- Wrote Adobe Flash language games for American Indian educational curricula
- Imported/exported lexicons between XML, FoxPro, Shoebox/Toolbox database formats

Nuance Communications, Burlington, Massachusetts   2007- 2008
Hungarian Language Consultant
- Corrected pronunciation lexicon using manual and rule-based strategies
- Compiled additional Hungarian language resources to suggest out-of-vocabulary lexical items

Indiana University, Bloomington, IN   1999-2005
Associate Instructor: Mathematics, Linguistics, and Psychology Departments
- Courses taught: Algebra, Finite Math, Precalculus, Calculus, Statistics, and Intro. Linguistics

Dragon Systems (Lernout & Hauspie), Burlington/Newton, Massachusetts   2001
Language Model Engineer in medical and research divisions
- Ran expectation-maximization and related proprietary tools to build hospital-specific language models for ASR to speed transcription of medical reports
- Curated and managed medical data by automatically sanitizing sensitive patient information preprocessing prior to tokenization and language modeling.
- Rewrote in-house language modeling tools from Perl into C for performance optimization

EDUCATION	**Indiana University**, Bloomington, IN; Doctor of Philosophy, Linguistics Dissertation: Quantitative studies in Hungarian phonotactics	2010 GPA 3.8/4.0
	Indiana University, Bloomington, IN; Master of Arts, Computational Linguistics Studied natural language processing and computational phonology	2003 GPA 3.8/4.0
	Indiana University, Bloomington, IN; Master of Arts, Mathematics Specialized in model theory, recursion theory, and algebra	2000 GPA: 3.7/4.0
	Bucknell University, Lewisburg, PA; Bachelor of Science, Magna Cum Laude Mathematics major, Computer Science minor	1999 GPA: 3.8/4.0

ADDITIONAL EDUCATION	**The Wharton School, University of Pennsylvania** Executive Education Certificate in Business Essentials / Management. Semester courses in Management, Human Resource Management, Business Strategy, Negotiations and Conflict Resolution, and Marketing.	2012
	Linguistic Society of America, Linguistics Summer Institute, Stanford University Intensive coursework in NLP and industrial linguistics. Attended courses taught by Jurafsky, Manning, Knight, Koehn, Bird, and Sproat.	Summer 2007
	Budapest Institute of Technology, Budapest, Hungary Budapest Semesters in Mathematics elite study abroad program	Winter-Spring 1998

TECHNICAL SKILLS

Languages: Ruby, Python, R, Perl, C, Javascript, ActionScript, Java
Code Libraries: Qt, CodeIgniter, Rails, JQuery
Tools: MySQL, mongo, git, Praat, Audacity, Matlab, Mathematica, SPSS

NATURAL LANGUAGES

Fluent: Hungarian, English (native)
Reading proficiency: German, Spanish
Structural knowledge: Turkish, Meskwaki, Lakota, Pawnee, Lusaamia (Bantu)

LEADERSHIP

Software Development team lead, Linguistic Data Consortium, 2012-present
Ultimate Frisbee team captain, various league teams, 2009-present
Bloomington Community Bicycle Project, Volunteer and Outreach Coordinator, 2000-2008
Hungarian Cultural Association, Officer 2002-2007, President 2006-2007
Indiana University Linguistics Club Working Papers, Assoc. Editor 2002, Chief Editor, 2003-2004
Mathematical Association of America, Bucknell Chapter, President 1997

HONORS

Foreign Language Area Study Fellowship, 2003-2004
Graduate Assistantship in Areas of National Need for Mathematics, 1999-2000
Barry M. Goldwater Fellowship, 1998-1999; one of eight Bucknell students all-time to receive honor
Phi Beta Kappa, 1998; Pi Mu Epsilon (mathematics honorary) 1998
Boy Scouts of America Eagle Scout, 1995
American High School Mathematics Exam, State of Ohio, 22nd place overall, 1994